DOCUMENT RESUME

ED 069 055 EC 050 105

AUTHOR	Hillman, Stephen B.
TITLE	The Effects of Question Type and Position on Four Types of Learning Among Mentally Retarded Children. Final Report.
INSTITUTION	Indiana Univ., Bloomington. Center for Innovation in Teaching the Handicapped.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE	Jul 72
GRANT	OEG-9-242178-4149-032
NOTE	135p.; Final Report 17.3
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Educable Mentally Handicapped; *Exceptional Child Research; *Inquiry Training; *Learning Processes; Logical Thinking; *Memorizing; Mentally Handicapped; Story Telling; Teaching Methods; Time Factors (Learning)

ABSTRACT

Examined were the effects of memory and inference questions on learning in 90 educable mentally retarded children (aged 9 to 13 years) who were read a short story, each section of which was preceded or followed by a training question which cued relevant information. Results indicated that those questions which followed the presentation of sections of the story were more effective in increasing short term achievement than were those which preceded the sections of the story, that close temporal proximity between critical information to be learned and the question led to greater learning, and that those subjects who received the same type of training question and criterion question performed better on the delayed relevant criterion tests (after 15 minutes) than did those who had a mixture of question types. No differences were found between groups on the incidental criterion tests. Appendixes included the short story, the criterion tests, and the training questions. (Author/GW)

ED 069055

Center for Innovation in Teaching the Handicapped

School of Education, Indiana University, Bloomington

THE EFFECTS OF QUESTION TYPE AND POSITION ON FOUR TYPES OF LEARNING AMONG MENTALLY RETARDED CHILDREN

STEPHEN B. HILLMAN

ED 069055

THE EFFECTS OF QUESTION TYPE AND POSITION ON FOUR TYPES OF LEARNING AMONG MENTALLY RETARDED CHILDREN[1]

Stephen B. Hillman

July, 1972

Final Report 17.3

Center for Innovation in Teaching the Handicapped

Indiana University

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

[1]This research was supported by grant #OEG 9-242178-4149-032 from the U. S. Office of Education, Bureau of Education for the Handicapped to the Center for Innovation in Teaching the Handicapped. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

EC 050 105

THE EFFECT OF QUESTION TYPE AND POSITION ON FOUR TYPES OF LEARNING AMONG MENTALLY RETARDED CHILDREN

Stephen B. Hillman

Center for Innovation in Teaching the Handicapped

Indiana University

Abstract

That questions play an important role in the daily instructional activities of teachers is supported by a considerable amount of descriptive research. However, the effects of these questions on student learning has received little attention. The purpose of the present study was to examine the effects of questions on learning among mentally handicapped children. The four types of learning were identified as relevant remembering, incidental remembering, relevant inferring, and incidental inferring.

Educable mentally retarded (EMR) children were asked to listen to a short story which was broken down into 10 sections. Each section was either preceded or followed by a training question, the purpose of which was to cue the relevant information. There were two types of questions, remembering and inferring. Any given subject received only one type. After completion of the story there was a 15 minute rest period during which the children played with stick figures. This was followed by a 20-item free-recall criterion test of two item types--remembering and inferring, a given subject receiving only one type of item. This did not have to be of the same type as the training question, and indeed, in half the cases the training questions and criterion questions were of different types.

Results of this study indicated that those questions which followed the presentation of sections of the story were more effective in increasing short-term achievement than were those which preceded the sections of the story; that close temporal proximity between critical information to be learned and the question lead to greater learning; and that those subjects who received the same type of training question and criterion question performed better on the delayed relevant criterion tests (after 15 minutes) than did those who had a mixture of question types. No differences were found between groups on the incidental criterion tests.

The results are discussed in terms of interference and memory theory and implications for the teaching of mentally handicapped children were identified.

PREFACE[1]

One of the Center's continuing goals has been to study the effects of teacher interactive behavior on the learning of mentally handicapped pupils and to invent ways in which teachers may be better trained in an attempt to increase pupil achievement.

The research study to be reported herein is part of a much larger attempt to reach this goal by examining the effects of teacher cognitive demands on pupil learning. Considerable time and effort have gone into the development of the Individual Cognitive Demand Schedule (Lynch and Ames, 1971b). This observation system systematically records the nature of the cognitive activity between teachers and students. The instrument has been used widely in descriptive studies of both normal and special classes. While considerable correlational evidence and theory exist to indicate the importance of teacher cognitive demands, little experimental data can be found to support the validity of statements concerning the differential effects of different types of questions. The study which follows is one of a number which attempt to shed light on this important question.

---

[1]The author wishes to thank Drs. William W. Lynch, Melvyn I. Semmel, Richard L. Turner and Robert B. Cairns for their helpful suggestions throughout the development and execution of this project.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	ix
Chapter	
I. INTRODUCTION	1
Statement of the Problem	1
Review of Research	2
Classroom Questions	2
Questions and Mathemagenic Behaviors	8
Relevant Characteristics of Educable Mentally Retarded Children	15
Objectives of the Present Study	20
Description of Learning Task	21
Definitions	23
Hypotheses	23
II. METHOD AND PROCEDURES	26
Design of the Study	26
Subjects	27
Materials	28
The Story	28
Rest Activity (Cuisenaire Rods)	30
Criterion Test	31
Experimenters	31
Procedures	32
Scoring of Tests	34
Analysis of Data	35
III. RESULTS AND INTERPRETATION	38
Between Trials Results and Interpretation	38
Delayed Criterion Test Results and Interpretation	53
Results and Interpretation with Respect to Mathemagenic Behaviors	74

Chapter Page

IV. DISCUSSION. 84

Discussion of Results 84
Implications for Teaching the Mentally Handicapped. . . . 94

V. SUMMARY 97

REFERENCES. 102

APPENDICES. 109

Appendix A: The Story. 109
Appendix B: Criterion Tests. 114
Appendix C: Training Questions 116
Appendix D: Instructions to Experimenters. 119
Appendix E: Item Analysis. 122
Appendix F: Correlation Matrix 124

LIST OF TABLES

Table		Page
1.	Experimental Design	22
2.	Means and Standard Deviations of Between Trials Scores	39
3.	Analysis of Variance of Between Trials Total Scores	40
4.	Stepwise Multiple Regression Analysis on Close Proximity Scores	47
5.	Stepwise Multiple Regression Analysis on Distant Proximity Scores	48
6.	Stepwise Multiple Regression Analysis on Total Between Trials Score	49
7.	Means and Standard Deviations of Relevant Delayed Criterion Test Scores (Based on a possible total of 20 points)	55
8.	Analysis of Variance of Relevant Delayed Criterion Test Scores without Control Group	56
9.	Analysis of Variance of Relevant Delayed Criterion Test Scores with Control Group	59
10.	Means and Standard Deviations of Incidental Delayed Criterion Test Scores (Based on possible total of 20 points)	61
11.	Analysis of Variance of Relevant and Incidental Delayed Criterion Subtest Scores with Control Groups	62
12.	Analysis of Variance of Total Delayed Criterion Test Scores without Control Group	68
13.	Stepwise Multiple Regression Analysis on Relevant Delayed Criterion Test Scores	71
14.	Stepwise Multiple Regression Analysis on Incidental Delayed Criterion Test Scores	72
15.	Stepwise Multiple Regression Analysis on Total Delayed Criterion Test Scores	73

Table		Page
16.	Analysis of Variance of Incidental Delayed Criterion Test Score without Control Group	76
17.	Analysis of Variance of Incidental Delayed Criterion Test Scores with Control Group	78
18.	Serial Analysis of Variance of Between Trials Scores. . .	79

LIST OF FIGURES

Figure		Page
1.	Graph of means of the interaction between location of information in sections of the story and position of question.	42
2.	Graph of the mean scores involved in the interaction of question type x question position x experimenter (T x P x E).	44
3.	Graph of the mean scores involved in the interaction of training question type and criterion type (T x C) -- Relevant Score.	57
4.	Graph of the interaction between criterion type and groups (C x G).	64
5.	Graph of interaction between groups and subtests on delayed criterion test (G x RI)	65
6.	Graph of the mean scores involved in the interaction of training question type and criterion type on total criterion score	69
7.	Graph of means of serial analysis by trials for pre- and post-question conditions.	80
8.	Graph of means of serial analysis by trials for remembering pre- and post-question conditions	81
9.	Graph of means of serial analysis by trials for inferring pre- and post-questions conditions.	82

Chapter I INTRODUCTION

Statement of the Problem

Children classified as being mentally retarded are characterized by slow intellectual development, poor performance on intellectual tasks, and poorer ability in self-guided learning than their normal IQ counterparts.

Different theories have been proposed to explain these observed inadequacies. Zigler (1966) focused on motivational factors. Others hypothesize deficits in attention (Semmel, 1965; Zeaman and House, 1963), short term memory (Ellis, 1963, 1970) and in organization of stimulus input (Spitz, 1966).

Denny (1964) hypothesized that the performance of mentally retarded children may be due to a deficit in incidental learning. The research of House and Zeaman (1963) tends to support this view in demonstrating that retarded children do not appear to attend to relevant cues, but that learning does progress once attention is directed to the appropriate cues. Other research (Semmel and Williams, 1968) has shown mentally retarded children to be inferior in both intentional and incidental learning, but that retarded children can achieve normal levels of incidental learning if given sufficient exposure to the stimulus material. Thus it appears that retarded children may require special attention and direction in helping them to achieve certain knowledge and learning skills which come more easily to normal children.

Several studies have shown that educable mentally retarded (EMR) children are capable of productive thinking (Cawley and Chase, 1967;

Prehm and Crosson, 1969; Rouse, 1965; Smith, 1967; Tisdall, 1962). Yet other research (Brophy and Good, 1970; Deutch, 1966; Lynch and Ames, 1971) indicates that low-ability children and children for whom teachers have low expectations miss out on opportunities for intellectual stimulation from the teacher by having fewer questions directed at them and that teachers frequently call on students they expect to give the right answer. Still, it remains to be demonstrated that questions are capable of being effective stimuli in increasing the achievement of retarded children. It was the attempt of the research to be reported here to study the effects of different types of questions on both intentional and incidental learning for the purpose of increasing the achievement of educable mentally retarded children.

Review of Research

The research pertinent to the hypotheses of this investigation are presented in three separate sections. The first concerns itself with studies of classroom questions, the second section with studies using questions to produce mathemagenic behaviors, and the third section with characteristics of educable mentally retarded children germane to this study.

Classroom Questions

That questions play an important role in the daily instructional activities of teachers is supported by at least 60 years of research. Stevens (1912) estimated that 80% of the time spent in school was occupied by question-and-answer recitations. Floyd (1960) found an average of 384 questions being asked by primary-grade teachers in a typical day;

Moyer (1966) found an average number of 180 questions per elementary science lesson, and Schrieber (1967) found that fifth grade teachers asked an average of 128 questions per social studies lesson. A recent study (Lynch & Ames, 1971a) demonstrated that there were no significant differences with respect to the number of questions asked per hour of instructional time between normal and special education classes.

Descriptive studies concerned with the type of question asked in verbal discourse have demonstrated that approximately two-thirds of the questions asked require direct recall of textbook information. Stevens (1912) was the first to describe this phenomenon. More recently Floyd (1960), Guszak (1967) and Schrieber (1967) have demonstrated similar percentages in elementary school situations; Gallagher (1965) and Davis and Tinsley (1967) have observed similar percentages in high schools. In fact these and other studies prompted Gall (1970) to conclude that teachers' questioning practices appear to be unchanged over this lengthy period.

Results of a recent study on questions in normal and special education classrooms (Lynch & Ames, 1971a) showed that teachers in each type of class seem to demonstrate similar percentages of higher and lower questions asked during instruction. Two other studies (Fine, Allen, & Medvene, 1968; Minskoff, 1967) show that teachers of elementary mentally retarded children use the greatest percentage of factual questions with teachers of normal elementary children using fewer and teachers of gifted high school children the fewest.

While educators have been concerned with the type and frequency of questions asked in classrooms for a long time, and more recently with questions in special education classes, little research has been done on the effects of different types of questions. The few studies which have examined the effects of questions are reported here. The subjects for these studies were all normal children or college students. The author was unable to locate any studies using educable mentally retarded children as subjects.

A study done by Hilda Taba (1966) sought to clarify the effects of teaching strategies on the cognitive functioning of elementary school children. Several findings from this study are of importance here. First, in both experimental and control groups, teachers were relatively successful in getting students to give the response they sought (a finding later supported in a study by Lynch and Ames, 1972). Second, although not consistent among all experimental classes, evidence from tests developed for this study showed that in ability to discriminate, to infer from data, and to apply known principles to new problems, the students of teachers who had been trained in the skills of the three cognitive tasks were superior to those in classes with untrained teachers. And thirdly, the results showed that the use of specific teaching strategies designed to foster development of cognitive skills seemed to make a difference in the general productivity of thought in terms of both higher levels and complexity of thought. "The most important observation that can be made from the data," Taba states, "is the centrality and power of the teacher's role in initiating cognitive operations and determining

which kinds are open to students" (1966, p. 228). The fact that students generally gave what teachers sought indicates the power that a questioning strategy has in determining those cognitive operations in which students engage.

In examining the relationship between the variables of question type and student achievement, Hunkins (1967, 1968) had one group of sixth grade subjects work exclusively with knowledge-level questions, while a second group worked with analysis-evaluation questions. Results showed that the analysis-evaluation group earned a significantly higher score on a specially-constructed posttest than did students who answered questions that stressed knowledge. While Hunkins' findings are of interest, they must be viewed with caution because of serious methodological considerations. Questions may also be raised about the definitions of analysis and evaluation used in this study.

Wright and Nuthall (1970) reported a study in which they explored the relationship between teacher behaviors and pupil achievement. Teacher behavior variables were identified from tape recordings and correlated with achievement test scores (developed on the basis of the lesson content outline especially for this study) which had been corrected for pupil intelligence and prior knowledge. Among other findings, Wright and Nuthall reported that the mean class achievement scores correlated significantly with patterns and kinds of teacher questioning. Some teachers tended to ask one question at a time, while others frequently asked two or more in rapid succession in a single utterance. The data showed (a) that the tendency to ask one question at a time was positively related to achieve-

ment, (b) the tendency to ask several questions was negatively correlated with achievement, and (c) the greater the percentage of a teacher's questions which were closed (i.e., required single statements of fact, description, definition, naming) as opposed to open (i.e., required statements of opinion, evaluation, explanation, inference), the higher the achievement of the pupils.

A fifth study having pupil cognitive learning as a dependent variable was conducted by Ladd and Anderson (1970) in which they investigated the effects of the level of inquiry of teachers' questions on the achievement of 1000 ninth-grade earth science students in 40 classes. A median split was used in separating the 40 participating teachers into equal size groups of low- and high-inquiry teachers based upon observations of their teaching behavior. Results of this study, with adjustments made for intelligence, show that the students of high-inquiry teachers performed significantly better on tests which contained (a) low inquiry questions only, (b) high inquiry questions, and (c) both high and low inquiry questions. The between-group differences were significant beyond the .001 level. The authors thus concluded that "teachers' questioning behavior strongly influences student achievement." (p. 398)

Two investigations (Furst, 1967; Thompson & Bowers, 1968), studying the cognitive level of classroom discourse, found a positive relationship between the amount of cognitive variation employed by the teacher during classroom discourse and student criterion performance.

In a study by Chall and Feldman (1966) teachers whom observers rated as emphasizing the stimulation of thought and skills tended to have

higher-achieving classes on reading subtest scores such as word reading, paragraph meaning, spelling and word study than their information and skill counterparts. Data from this study also indicated a positive relationship between student achievement and a rating of whether the cognitive level of the teacher's lesson appeared to be "just right most of the time."

Several studies have been conducted in which questions were classified into two types. The results of seven of these investigations are reported here. Of these seven, significant results were not obtained in three. In two of these studies (Harris and Serwer, 1966; Harris et al., 1968) attempts were made to correlate the OScAR-R subscales with reading achievement among disadvantaged children. In the third study Perkins (1965) attempted to correlate teacher behavior with lack of achievement among high ability students. A fourth study by Wright and Nuthall (1970) found open and closed questions unrelated to achievement when the frequencies were converted to percentages, a significant relationship was obtained with the closed questions, showing a positive relationship to pupil achievement. Of the three remaining studies, all reported significant results. Kleinman (1964) reported that teachers with high-achieving junior high school students asked more "high-level" questions during their science lessons; Spaulding (1965) found "open-ended" questions to have a negative correlation with achievement among elementary students; and Thompson and Bowers (1968) reported that high achieving students had teachers who asked a mixture of convergent and divergent questions.

Rosenshine and Furst (1971) report that they had located two studies which used multiple classifications of teacher questions. Both of these

studies (Conners & Eisenberg, 1966; Solomon, Bezdek, & Rosenberg, 1963) had significant results. The Conners and Eisenberg (1966) study in particular demonstrated an important and strong relationship between intellectual activities of the teacher and increased pupil scores on the Peabody Picture Vocabulary Test (PPVT).

The general conclusion to be drawn from these studies seems to be that there is at least a correlational relationship between teacher questioning behaviors and student achievement in normal children. Exactly what that relationship is remains unclear. The correlational evidence has indicated several profitable variables for more carefully controlled experimental studies. Some of these variables which appear to have a positive correlation with achievement are: (a) the tendency to ask one question at a time as opposed to several at once (Wright & Nuthall, 1970); (b) the amount of cognitive variation employed by the teacher (Furst, 1967; Thompson & Bowers, 1968); (c) the opportunity the students have to learn at the cognitive level appropriate to the criterion achievement test (Chall & Feldman, 1966; Hunkins, 1967, 1968; Ladd & Anderson, 1970; Taba, 1966); and (d) the tendency to ask "closed" as opposed to "open" questions (Spaulding, 1965; Wright & Nuthall, 1970).

In summary, most of the research on the effects of teacher questions has been of a correlational nature. The research has produced a number of interesting correlates with student achievement, but the validity of these correlates needs to be further explored in experimental studies.

Questions and Mathemagenic Behaviors

In recent years a large number of studies have been conducted with respect to cognitive learning from written materials. E. Z. Rothkopf coined

the term "mathemagenic" to refer to "those student activities that are relevant to the achievement of specified instructional objectives in specified situations or places" (Rothkopf, 1970). Rothkopf assumes that these student activities are modifiable, that if improperly controlled or uncontrolled may lead to irrelevant learning and that the learner adapts his activities to the requirements of training questions or orienting tasks (1970). In attempting to gather empirical evidence in support of these assumptions, Rothkopf and others have focused mainly on the influence of training questions on learning from written materials. The following selected studies will provide the reader with a general review of the nature of this work.

Rothkopf (1966) investigated the effects of adjunct, test-like questions on learning from written materials. In this study he asked college students to read a 5,200 word selection from Rachel Carson's The Sea Around Us which was broken down into seven sections of approximately equal length. As criteria he used a test composed of 14 questions intended to measure specific learning resulting from experimental questions asked in the text and a test of 25 items not used as experimental questions, called a general test. There were six treatment groups: (1) "SBA" - subjects were given two questions shortly before each of the seven sections. After writing his guess, the subject was given the correct answer; (2) "SB" - same treatment as SBA except that the subject was not provided with the correct answer after he made his guess; (3) "LBA" - subjects were given all 14 questions at once just before starting to read the chapter. They were given the correct answer after they made a guess on each question; (4) "SAA" -

the subjects were given two questions immediately after each of the seven sections. The correct answer was provided as soon as the subject responded to each question; (5) "SA" - same treatment as SAA except that the correct answers were not provided after the subject gave his answer; (6) Control - no questions were given in the text. A direction reference group (DRG) was added to evaluate the effects of questions compared to care-inducing directions.

The results of the study indicated that the SA, SAA, and DRG groups performed better on the 25-item general test than all other groups. All groups which received questions performed significantly better on the 14 item specific learning test than those groups which did not. Those groups receiving answers after responding did better on the 14 item specific test than did those with no feedback. It was evident from these results that (1) asking questions after reading the relevant test passages facilitated both specific and general learning, (2) questions presented before the relevant text passage produced only question-specific facilitative effects, and (3) that question-specific-effects were most noticeable when the correct answer had been given to the subjects after they responded. Rothkopf concluded from this that adjunct questions, unlike specific directions, may shape effective inspection behavior and are also useful in teaching specific skills.

A study by Rothkopf and Bisbicos (1967) hypothesized selective facilitative effects of interspersed questions on the learning of written materials. The 252 high school subjects were asked to read a 36-page, 9000-word passage about animals and minerals found in the sea. Two questions appeared in the text per each three-page zone, but the questions differed

in location (before or after the relevant segment) and in required response. Different treatment conditions saw questions restricted to one of the following response types (a) either a quantitative term or name, (b) a common English or a technical word, (c) a mixture of (a) and (b). Each treatment group responded to the same 48-item criterion test after having read the passage. Results indicated that learning of the several categories of text content was facilitated by appropriate questions seen immediately after exposure to the relevant text segment as opposed to those seen before.

In a somewhat related study using a similar experimental design, Frase (1968b) was able to replicate the finding that retention was highest when questions were placed after the appropriate material. The 128 college students in this experiment were asked to read a 2000-word passage concerning the life of William James. Instead of questions being placed at two or three page intervals, however, questions were paced at the rate of one every 10, 20, 40, or 50 sentences. The data indicated that retention increased with the frequency of posttreatment questions, but it decreased with frequent pretreatment questions. Question mode (multiple choice or constructed response) in terms of questions appearing in the text was also a variable, but had no effect.

The problem explored in the next study (Frase, 1968a) was to determine what happens to the retention of information contained in a passage when an orienting question is asked which requires the processing of a relatively large or small amount of the total information contained in that particular passage. Eighty-four college subjects were allowed 20 seconds to read a question and a 36-word paragraph. While the paragraph was the same for all subjects, the questions differed. One group of subjects read

a specific question, another group a comparative question, and the third group a general question. Each question was read by each experimental group before reading the paragraph. The results of the study showed that (a) the most precise question (i.e., specific > comparative > general) led to the most efficient acquisition of the specific stimulus-response association (i.e., more subjects in the specific question group passed the test item which was relevant to their question), and (b) when performance on the total retention test was the criterion, the groups scored in the same order (i.e., specific > comparative > general). While result (a) is consistent with experimenter's hypothesis, result (b) is just the opposite of what was predicted. It was reasoned that general orienting questions would require the subjects to process greater amounts of information and thus their general retention would be higher. Though the results did not support this positon, they did supply evidence for the selective information rejection (attention) position suggested by Berlyne (1965) and Schroder, Driver, and Streufert (1967) in which it is hypothesized that the greater uncertainty created by the comparative and general questions forced the subjects to engage in information rejection strategies in order to reduce the information load of the paragraph. Data derived in another project as part of this same study indeed add support to this position. To quote Frase, 'The general conclusion seems to be that as effective uncertainity or information load increases, precise control over reading behavior becomes more imperative" (p. 201).

Another series of studies conducted by Frase (1969) and reported in monograph form induced subjects to think about text material by having them deduce conclusions from that material. The conceptual characteristics

of the text material were analyzed in order to permit predictions about which text items would enter memory as a function of different orienting directions. The ability to control learning from text material, it was stated, hinges upon an adequate understanding of this interactive process. It was clear from the results of the first experiment in this monograph that when a certain text item was a component of a problem solution, it had access to memory and was higher in recall than if it was not part of the problem solution. The basic hypothesis here is that, while subjects might scan an entire passage for the information necessary to draw a certain inference as communicated by an orienting direction or question, the text which is not relevant to that conclusion will receive only minimal processing and not have access to memory. This finding held for all three of the experiments in the monograph. Specifically though, experiment #1 showed that the recall of text items which mediated problem solution was greater than for those text points which did not mediate problem solution; and experiments #2 and #3 demonstrated that inducing higher levels of information processing adds new items to memory and thus raises the over-all level of recall, but does not increase the number of correct inferences.

A final study in this section concerns itself with the effects of written versus orally-communicated questions on learning from written materials (Rothkopf & Bloom, 1970). Sixty-three high school students studied a 16,000-word earth science text which was presented to them individually on 180 slides. In one experimental group, a written question related to the previous reading appeared after every sixth slide. The subject then wrote down his response on a piece of paper, but received no feedback. The second experimental group received an oral question asked by a teacher after every sixth slide. The

subject gave his response, but received no feedback. The control group received no questions. The results indicated that the oral question group scored significantly higher on a recall criterion test than did the written question group, and that both groups scored significanlty higher than the control.

In summarizing briefly, it seems clear that different questions can in fact produce different learning outcomes (e.g., Frase, 1968a, 1968b; Rothkopf and Bisbicos, 1967). It also seems clear that when learning from text material, questioning can produce more effective learning than no questioning (Rothkopf, 1966; Rothkopf & Bisbicos, 1967). Questions which are asked after a subject reads a particular portion of the text as opposed to questions which are asked in advance of his reading also produce better learning (Frase, 1967, 1968a,b; Rothkopf, 1966; Rothkopf & Bisbicos, 1967). The most interesting result of the studies on adjunct questions, according to Rothkopf (1970), is that they demonstrate that mathemagenic activities are adaptive and that "the shaping of mathemagenic activities in an instructional fashion by environmental events (or contingencies) is a practical possibility (p. 333)."

Unfortunately all of these studies have been done on college or high school students and it remains to be seen whether or not the behaviors of mentally handicapped children are modifiable under similar stimulus control and during a variety of learning tasks. Because of the consistency of the reported pre-post question results in favor of the post question in producing greater learning, this would seem to be an important variable in testing the adaptability of the mentally handicapped to different question stimuli.

Relevant Characteristics of Educable Mentally Retarded Children

Mentally retarded children tend to be slow to develop intellectually (Robinson & Robinson, 1965) and to do poorly on intellectual tasks in general (Denny, 1964). Several attempts to explain this have postulated various types of deficits characteristic of the educable mentally retarded (EMR) child (e.g., deficits in attention to appropriate cues (Semmel, 1968; Zeaman and House, 1963); short-term memory (Ellis, 1963, 1970); organizational strategies (Spitz, 1966); and in the development of learning sets Kaufman and Prehm, 1966). The following represents a brief review of theory and research dealing with each of these orientations.

Zeaman and House (1963) have argued that the observed learning deficit of retarded children may be accounted for by a lack of attention. Stated more precisely, they theorize that (1) attention is limited to only one (or at most, a few) of the many possible stimulus dimensions available to the subject at any particular moment, (2) subjects may learn to attend to or disregard stimulus aspects as a function of differential reinforcement, and (3) cues for instrumental learning are those aspects of the stimulus which are being attended to (p. 212). Zeaman and House (1963) go on to state that

> If our analysis of retardation and attention is correct, the secret of successful training of moderately retarded children lies in the engineering of their attention. . . . one should seek ways of increasing the attention value of the relevant cues. (p. 218)

Drawing on the work of Zeaman and House (1963), Denny (1964) hypothesized that the performance of mentally retarded children may be due to a deficit in incidental learning, and contended that the basic attentional

problem hypothesized by Zeaman and House (1963) may result in poor incidental learning. Several experimenters have reported results consistent with this hypothesis (e.g., Goldstein and Kass, 1961; Semmel and Williams, 1968). Thus it appears that retarded children may require special direction and additional help in school learning situations.

Research on short-term (STM) and long-term memory (LTM) involves considerable definitional problems. Distinctions between these two concepts are often vague and arbitrary. Nonetheless, several authors have proposed theoretical distinctions between STM and LTM (Ellis , 1963, 1970; Hebb, 1949; Peterson, 1966). The work of N. R. Ellis is discussed here because it deals focally with the issue of memory and the retarded child.

In 1963 Ellis hypothesized that the inadequate behavior of retardates was in part due to a stimulus trace diminished in both strength and duration. Further, he predicted that (1) when the performance of retardates and normals were compared on tasks which required the bridging of a time gap, that the retardate's behavior would suffer in comparison; (2) as the magnitude of the temporal separation between events increased, the retardate's performance would deteriorate even more; and (3) that the stimulus trace would show a developmental trend with both strength and duration increasing as a function of age and intelligence. In theorizing about the role of memory in learning, Ellis predicted poorer performance by retardates as compared with normals on a wide range of tasks such as reaction time, delayed response, paired-associate learning as well as simple retention.

Results of experiments designed to test the above theory led Ellis (1970) to hypothesize a multi-process conception for retention of supraspan

information. This reformulation postulated the presence of two memory processes -- primary memory (PM) and secondary memory (SM). Research (Ellis & Hope, 1968) had shown the recency and primacy segments of the serial position curve to be discontinuous processes. Primacy could be influenced by rate of presentation, or by delay prior to recall. It was hypothesized that rehearsal strategies may be responsible for facilitating primacy performance. Rehearsal appeared to have no effect upon recency.

Research results (Ellis, 1970) suggest that primary memory (PM) in the retardate and normal may not differ. But, while this recency performance is comparable in both retarded and normal subjects, adult and adolescent retardates display a lower primacy performance than normal subjects (Ellis, 1969). As is the case with normal subjects, simple rehearsal or labeling facilitates recency and depresses primacy effects in a probe-type memory task (Ellis, 1969). From these findings Ellis (1970) concluded that while normal subjects showed evidence of using both primary and secondary memory, in retardates the secondary memory process fails to function in the normal fashion. Ellis further concluded that active rehearsal strategies appeared necessary for secondary memory but not for primary memory and that the retardate's deficiency may be due to a failure of the rehearsal mechanism(s). Belmont and Butterfield (1969) argue in a similar fashion by suggesting that the short-term memory deficit of retardates results from and acquisiton deficit which is probably due to a failure to actively rehearse stimulus input after it enters primary memory rather than due to defective retention or retrieval. It seems plausible, Sitko and Semmel (1971) argue, to hypothesize a deficit in secondary memory which is determined by "secondary organization rather than a deficit in primary or short-term memory."

Several investigators have argued that the need for individuals to store a vast amount of information in a limited storage capacity requires the individual to impose some form of information reduction on the stimulus input (Bruner, Goodnow and Austin, 1956; Bruner, Oliver and Greenfield, 1966; Mandler, 1967; Miller, 1956). These views which characterize the child as an "active" learner deal primarily with the ability of an individual in selecting, processing, storing, and retrieving relevant information. Human memory, for these theorists, is essentially an active process of the organization of stimulus input by the learner.

Spitz (1966) has suggested that the slower learning performance of mentally retarded children on cognitive and memory tasks may be due to the faulty or inefficient organization of stimulus material to be learned. Spitz maintains, however, that "the question is not whether or not retardates group or organize materials, but rather under what conditions, in what manner, and how efficiently they display this capacity" (p. 36).

Semmel (1967) in examining differences in cognitive organizational strategies used by EMR and nonretarded children suggests that two fundamental and qualitatively different strategies are involved. The first is called the sequential-associative strategy and the second is called the hierarchical strategy. EMR children, according to Semmel (1967), tend to use only the more primative sequential-associative strategy when processing language, while both the hierarchical and sequential-associative strategies are typically used by normal children. The sequential-associative strategy results in simple stimulus-response associations or sequentially dependent chains, while hierarchical processing results in the formation of a struc-

tured network of concepts, classes, systems and relationships and involves the abstraction of common attributes to form a generalized internal representation. Support for Semmel's positon has been demonstrated in several studies (e.g., Semmel and Bennett, 1970; Semmel, Barritt, and Bennett, 1970; Semmel, Barritt, Bennett, and Perfetti, 1968; Sitko, 1970).

One of the clear advantages of the learning set approach to the study of fundamental learning processes is that one is able to systematically record an index of learning efficiency, or in some cases skill development, over a series of problems or trials. The major credit for the development of learning set evaluation techniques belongs to Harlow for repeatedly demonstrating that organisms can learn a series of discrete discrimination problems with pregressively greater efficiency (1959).

Learning set research with mentally retarded subjects has produced the following results: (1) mildly retarded children tend to produce significantly more stimulus preservation errors than normal children of the same chronological age (Kaufman and Peterson, 1958, 1965); and (2) a relatively gross relationship exists between mental age or IQ and learning set formation (Ellis, 1958; Harter, 1965, 1967; Kaufman and Prehm, 1966; Stevenson and Swartz, 1958) with normal children doing better than the mentally retarded subjects.

Though the above research tends to show that mentally retarded children exhibit difficulty in acquiring learning sets, they also demonstrate that learning sets may be acquired by those same subjects. Indeed, some studies show that with appropriate procedures (e.g., overlearning of the original task) learning sets may be retained for as much as a year (Kaufman, 1971).

Objectives of the Present Study

School situations place certain demands upon the learner. One of the most obvious is that what is learned in school is transmitted verbally in classrooms. In the case of normal children much of this is through written materials, especially as the child gets older. But because of poor reading skills, the retarded child continues to learn predominantly through the processing of verbal information which is transmitted aurally. In view of the need which mentally retarded children have for more directed and structured learning activities, it is important that means be found to supply this structure and thus to facilitate their cognitive learning during listening tasks. Questions have been found to play a large role in the daily instructional activity of the classroom teacher, and indeed, some evidence suggests that questions may have positive effects on the learning of normal children. As of yet questions have not been demonstrated to have a significant influence on the behavior of mentally retarded children, though the cuing and structuring effects of questions offer considerable theoretical, as well as practical, appeal.

It was the intent of this study to demonstrate that questions have important theoretical and practical effects in enhancing the cognitive learning of educable mentally retarded children. This study attempted to answer the following general questions:

(1) Can questions be used effectively to facilitate the short term retention of information?

(2) Can questions be used to facilitate the longer term retention of information (after 15 minute delay)?

(3) Does the manipulation of question position relative to content material influence short-term retention? long-term retention (after 15 minutes)?

(4) Can questions be used to effectively influence information processing strategies?

(5) Can questions be effective in creating learning sets in mentally retarded children?

Description of Learning Task

The subjects in this study were asked to listen to a short mystery story of approximately 2400 words. The story was broken down into ten sections. In two of the experimental conditions subjects were asked one questions before each of the ten sections. In the other two experimental conditions subjects were each asked one question after hearing one of the ten sections. In all cases the subjects responded to the question after hearing the section pertaining to that question, but received no feedback to their responses. All subjects heard all ten sections and answered ten questions in total. A control group heard the story but received no questions during the listening task. Half of the subjects received remembering questions throughout the listening task. The remaining half received inferring questions.

A 20-item criterion test was administered after a 15 minute delay. Ten of the items were the same questions the subject had been asked during the story. These are called relevant items. The remaining 10 items the subject had not been asked yet and these are called incidental items. See Table 1 for a pictorial display of the this design.

Question Type

Question Position	REMEMBERING						INFERRING					
	E_1	E_2	E_3	E_4	E_5	E_6	E_1	E_2	E_3	E_4	E_5	E_6
Pre	8* ②	2 ②	5 ①	6 ②	16 ②	10 ②	49 ①	43 ①	48 ①	47 ①	40 ②	39 ①
	13 ①	14 ②	18 ②	17 ①	9 ①	3 ①	44 ②	50 ②	53 ①	54 ②	52 ②	51 ①
	1 ①	7 ①	12 ②	11 ①	15 ①	4 ②	38 ②	37 ①	41 ①	42 ②	45 ①	46 ②
Post	36 ②	24 ②	19 ①	32 ②	27 ①	21 ①	59 ①	72 ②	55 ①	56 ②	58 ②	69 ①
	23 ①	29 ①	26 ②	20 ②	34 ②	33 ①	66 ②	60 ②	68 ②	61 ①	63 ①	64 ②
	30 ②	35 ①	31 ①	25 ①	22 ②	28 ②	71 ①	65 ①	62 ②	67 ①	57 ①	70 ②
No	77 ①	83 ①	85 ①	79 ①	86 ②	87 ①						
	84 ②	78 ②	73 ①	88 ②	76 ②	75 ①						
	89 ①	90 ②	80 ②	74 ②	81 ①	82 ②						

Table 1. Experimental Design

* uncircled number is the subject number

① = relevant and incidental remembering criterion test

② = relevant and incidental inferring criterion test

E = experimenter

Definitions

1. RELEVANT QUESTIONS are defined as those questions which were asked by the experimenter during the listening activity.
2. INCIDENTAL QUESTIONS are those items which relate to the material covered during the listening activity (i.e., information contained in the story), but which had not been asked before.
3. REMEMBERING QUESTIONS are defined as those questions which require the child to recall something. It may be a straight forward summary or a piece of information which he is required to remember.
4. INFERRING QUESTIONS call upon the child to make interpretations or to draw conclusions from data. The child must arrive at his own interpretation, deduction, or conclusion from available information.
5. PRE-QUESTIONS are those questions which were asked in advance of the child's listening to a section of the story.
6. POST-QUESTIONS are those questions which were asked after the child had heard a section of the story.

Hypotheses

As a result of the findings of previous studies cited in the review of the literature, the following hypotheses were tested in this investigation:

On the analysis of between-trials scores (short-term retention)

1 - Children receiving pre-questions would perform better than children receiving post-questions.

2 - Children would score higher on trials where information necessary to answer the questions were in "close proximity" as opposed to "distant proximity" from the question.

3 - Children would perform better on trials where they received a pre-question and the information was close as compared with when they received a post-question and the information was close.

4 - Children who receive post-questions would show improvement over trials on between trials scores.

5 - Children who receive pre-questions would show no improvement over trials on between trials scores.

On analysis of delayed (after 15 minutes) criterion measures

6 - Children who received questions would score higher on the relevant criterion test which asked the same questions as were in the training than children who were in the control group and did not receive training questions.

7 - Children who received questions would perform better on the incidental criterion test than children who were in the control group and did not receive training questions.

8 - Children who receive pre-questions would perform better on the relevant criterion test than children who received post-questions.

9 - Children would perform better on the relevant subtest than on the incidental subtest.

10 - Children who received remembering training questions would score higher on the remembering relevant criterion test than children who had received inferring training questions.

11 - Children who received inferring training questions would perform better on the inferring relevant criterion test than children who had received remembering training questions.

12 - Children who received remembering training questions would score higher on the remembering incidental criterion test than children who had received inferring training questions.

13 - Children who received inferring training questions would perform better on the inferring incidental criterion test than children who had received remembering training questions.

14 - Total score criterion test results would be highest for children receiving the same type of criterion test as training question.

15 - Children who received post-training questions would perform better than children who received pre-training questions on the incidental criterion test where the items in the test were of the same type as the training questions.

CHAPTER II

METHOD AND PROCEDURES

The preceding chapter described the review of related literature and listed the objectives and specific hypotheses of the present investigation. This chapter describes the design of the study, subjects, materials, experimenters, procedures and analysis of the data.

Design of the Study

The study was designed as a 3x2x2x2x6 repeated measures design with the repeated measure being location of information (i.e., answer) within the section of the story. Question position was the first factor with three levels: pre-question, post-question, or no question (control group). The second factor was question type: remembering or inferring questions. The third factor, location of answer in the section, was the repeated measure and had two levels: first half of the section or second half of the section. Criterion type was the fourth factor and consisted of two levels: remembering or inferring questions. The fifth factor was experimenter and there were six levels.

There were four dependent variables: total between-trials score, relevant score after fifteen minute delay, incidental score after fifteen minute delay, and total score after fifteen minute delay.

A total of 8 experimental groups and 2 control groups were involved in this study. Each group had 9 subjects. Subjects in the experimental groups received questions either before or after listening to a section of the story. Each of the 10 sections of the story was accompanied by a different question but of the same type. All of the questions for any

particular subject were either pre-questions (before he heard a particular section of the story) or post-questions (subject heard the question after he heard the appropriate section of the story). In half of the sections of the story the answer to the questions was contained in the first half of the section. The second half of the section contained the information necessary to answer the question in the remainder of the sections. Within these limitations, the position of the answer in a particular section was assigned at random. The subjects in the control groups received no questions between sections of the listening task. In place of the question there was a period of twenty seconds of silence. Then the story continued.

After the fifteen minute delay-rest period where the subjects played with the Cuisenaire Rods, each subject was given a twenty item criterion test. This test could be of two types: remembering or inferring. Each subject received only one type of criterion test. Half of the subjects who received remembering questions during the listening activity received the remembering criterion test; the other half received an inferring criterion test. Similarly, half of the subjects who received inferring questions during the listening activity received the inferring criterion test; the other half received the remembering criterion test. One of the control groups received the remembering criterion test; the other the inferring. See Table 1 (p. 22) for a pictorial display of the above description.

Subjects

The subjects (N = 90) were mentally retarded children who were students in intermediate special education classes in Indianapolis, Indiana. They ranged in chronological age from 114 months to 168 months ($\bar{X}$ = 141.60,

S.D. = 10.90) with mental age ranging from 74.48 months to 132.88 months ($\bar{X}$ = 98.42, S.D. = 13.10). The sample included 54 males and 36 females. The schools from which these children were selected were chosen in such a way as to give as representative a sample of EMR children as the urban population from which they were drawn would allow. As such, schools were selected in order to give as complete a balance between race (i.e., black and white) and socio-economic status (i.e., lower class and middle class) as possible. A school was determined to be predominately black or white depending upon which race made up more than 50% of the population of that school and the socio-economic status of the school was determined by the Director of Research for the Indianapolis Public Schools. The children were sampled from a total of 11 different schools.

Subjects were randomly selected from special education classes within each of the 11 schools. A child was included in the study if he or she was one of those randomly selected and was not eliminated by that child's teacher because of either hearing problems, speech problems or severe behavior problems. Children were then randomly assigned to treatment conditions. A subsequent t-test of means confirmed that mental and chronological age were randomly distributed among treatment and control groups.

Materials

The Story

Listening to a short mystery story was selected as the learning task in this study. This was done for two reasons. First, was the importance given to learning about the effects of questions on the acquisition

of verbal information and the particularly great importance of learning through listening in real school situations. The second reason was that a short story would be more likely to keep the child on task while at the same time it could be broken down into 10 separate sections for the purposes of experimentation and analysis, with each section being treated as a separate unit.

The particular story the children were asked to listen to was a high interest-low vocabulary (vocabulary fourth grade) short mystery story of approximately 2400 words. It was broken down into ten distinct sections of approximately equal length. The story is about a boy named Leroy who finds a map which he believes marks the spot where bank robbers buried some stolen money. The spot, he finds out later, is in his backyard, but does not contain stolen money. Rather, it is the spot where some men did some digging in order to find out if there was oil under the ground there. The complete story may be found in Appendix A.

Each of the ten sections was re-written in such a way as to yield a total of four questions, two of which were remembering and two inferring. The operational definitions of these forms are derived from the work of Lynch and Ames (1971b). The most important distinguishing characteristic of remembering questions is that they call upon the child to recall something. He may be asked for a straightforward summary of something or a piece of information. It does not call upon the child to interpret or draw conclusions from data. Inferring questions are distinguished by the fact that they call upon the child to arrive at his own interpretation, deduction, or conclusion from available information. A task in this cate-

gory asks the child to "go beyond the data" and arrive at some sort of conclusion. The answer to one of the remembering questions and one of the inferring questions could be found in the first half of each section. The answer to the remaining remembering and inferring questions could be found in the second half of each section.

After the story was re-written, it was recorded on audio tape. Several professional story tellers made recordings and the one which was considered to be the best with respect to voice modulation, enthusiasm, word pronounciation, and general technical quality of the recording was finally selected. The story was put on tape and ultimately communicated to the experimental subjects through that medium in an attempt to reduce the variance which may have been accountable by the experimenters' reading speed, pronounciation, enthusiasm, etc.

Rest Activity (Cuisenaire Rods)

After listening to the story there was a rest period of fifteen minutes followed by the criterion test. During this fifteen minute period the experimenter engaged the child in an activity of playing with Cuisenaire Rods.[1]

Though the original intent of the developers of the Cuisenaire Rods was to help in the teaching of mathematics, the varying lengths and bright colors of the rods made them appropriate for making stick figures, pictures and three-dimensional objects. The intent was to use these mate-

---

[1]Cuisenaire Company of America, Inc.
9 Elm Avenue
Mt. Vernon, New York 10550

rials in a carefree, non-threatening and relaxing manner. They were not used instructionally in this study but only to provide a vehicle for filling the fifteen minute period of time with an enjoyable, but non-tiring activity.

Criterion Test

The criterion test was administered fifteen minutes after the listening activity. There were two types of criterion test: rememberin; and inferring. Each test was composed of twenty free recall type items in total, but each subject received only one type of test (i.e., either remembering or inferring but in no case both).

Each criterion test consisted of two subtests: a relevant subtest and an incidental subtest. The relevant questions were those which had been asked earlier during the listening task and as such had been heard by the subject once before. The incidental subtest contained items based upon the same story which the subject had listened to earlier, but were questions he had never heard before. All items were classified by type by two people who had demonstrated consistent reliability with the Lynch-Ames (1971) category system. Copies of the criterion tests may be found in Appendix B.

Experimenters

Six people served as experimenters in this study. Five were female; one was male. All were employed by the Center for Innovation in Teaching the Handicapped, Indiana University, Bloomington, Indiana, as either research assistants or associates and represented a variety of backgrounds. During a training session preceding the collection of the

data, the experimental procedures were rigidly standardized and practiced. Each experimenter received a copy of the experimental procedures. A copy of the instructions to experimenters may be found in Appendix D.

As a check, the author listened to the tapes of the experimental proceedings for all of the experimenters after the first day of data collection. Though few variations from the standard procedure were found, several minor infractions were noted and the appropriate experimenters were informed. The tapes were checked again after the second day of data collection and all errors in procedure were found to have been eliminated.

Procedures

Subjects were randomly selected from cooperating classrooms and taken one at a time to the experimental room by the experimenter. After establishing rapport with the subject, the experimenter indicated that they would be listening to a short mystery story which was played on a tape recorder. The experimenter introduced the story with the general statement "that the story is about a young boy who lives in a city and the problems he has when he tries to catch some bank robbers." He (she) was also told to expect a surprise ending. The experimenter stressed that it was important to listen very carefully as he was interested in how much they learned from the story and that they would be asked questions after the story was over.

During the listening activity the experimenter systematically introduced the questions. The questions were typed on a sheet of paper which the experimenter held before him. The experimenters had been in-

structed to read the questions as they were printed. These questions were of two types (remembering or inferring) and in one of two positions (pre or post). In the pre-remembering questions condition the experimenter asked a remembering question before the subject listened to each section. As in all other conditions, the subjects were instructed to respond to these questions orally after listening to each of the paragraphs. No feedback was given. The experimenter could only say Thank you or O.K., but could not indicate that the answer was right or wrong. The pre-inferring question group was asked one inferring questions before each paragraph was heard. In the post-question conditions the questions were asked one at a time but after each section had been heard. There were a total of ten paragraphs and thus, ten questions. The control groups received no questions but did listen to the same short story. Instead, they had a short twenty second break between each paragraph during which they just sat quietly. All of the questions were of the free recall variety and the experimenter wrote down all of the subjects' responses verbatim.

After the story was over, there was a fifteen minute rest period during which the subject and experimenter played with the Cuisenaire Rods. Each subject was informed before the 15 minute break that after they played for awhile, they would have a short test.

After the rest period the experimenter administered the twenty-item criterion test. Each of the questions was typed on a sheet of paper which the experimenter had before him. The subjects' responses were again

written down exactly as spoken. Each item could be read no more than two times in total. Feedback to these questions could be accepting (e.g, O.K., Thank you, Uh huh) or positive (e.g., good, fine, etc.). The objective here was to keep the subject on task, relaxed and responding.

All questions and answerswere tape recorded during the experimental activity.

After the test was over, the subject was thanked very much for his (her) cooperation and escorted back to the classroom. The subject was requested not to tell anyone about the story and that it would be a secret between them until the next day. Then he could talk about it if he wished.

Scoring of Tests

A random sample of answers which the experimenters wrote down were checked by two people independently, against the corresponding audio tape for accuracy of the written response. In as much as this showed the written response to be accurate more than 99% of the time, the scoring of the answers was done on the basis of the responses which were written down by the experimenters as opposed to the taped version of the subjects' responses.

The answers to each question were judged in turn; i.e., all of the responses to a given question were judged before the answers to another question. Each response was compared to that on the answer key. The answer key was made up in advance of any judging and was based upon the information contained in the story. Each response could receive one of

three possible scores. An answer which was completely correct earned 2 points; 1 point was given for half credit and 0 points for an incorrect response. All of the answerswere judged in such a way that the judge had no knowledge of the condition in which the responder to the question was. After all answers were judged, an estimate of reliability was obtained by re-scoring a random sample of answers. This produced 94.5% agreement between the first and second scoring. The results of an item analysis for each criterion test may be found in Appendix E.

Analysis of Data

The data were analyzed in the following manner in order to test the stated hypotheses:

(1) an item analysis was done on each of the two criterion tests

(2) all independent and dependent variable were intercorrelated (computer program BMD02D)

(3) the responses to the ten training questions used during the listening task were analyzed by a serial analysis of variance on trials data (4 x 10 repeated measures on trials; computer program ANOVAR by D. J. Veldman, 1967).

(4) analysis of the total between trials score was accomplished with a 2x2x2 ANOVA with repeated measures on the last factor (position of information in section of the story). The first two factors were question type and position (BMD08V)

(5) a 2x2x2x6 ANOVA was used to analyze the criterion data after the fifteen minute delay. This was done on the incidental and

total score dependent variables without the control group (computer program RCC5V).[1]

(6) a 2x2x6 ANOVA was used to analyze the relevant delayed criterion data. The criterion type was dropped from this analysis as type of question is accounted for in factor one and because these are relevant scores, the questions must always be of the same type (computer program RCC5V)

(7) a 2x5 ANOVA was used to analyze the relevant and incidental scores with the control group. The first factor was criterion type and the second factor was groups (computer program RCC5V)

(8) the relevant criterion measures without the control were compared in a 2x2x6x2 ANOVA (computer program RCC5V).

(9) the relevant criterion measures with the control group added were compared in a 2x5 ANOVA (computer program RCC5V).

(10) the relevant and incident criterion measures were compared in a 2x5x2 ANOVA (computer program BMD08V).

(11) a Stepwise Multiple Regression Analysis (Nie et al., 1970) was done on each dependent variable. Each analysis yielded an equation in the following form:

$$X_0 = b_1X_1 + b_2X_2 + \ldots + b_nX_n + R$$

where

X_0 = dependent variable score

b_i = unnormalized coefficient

X_i = independent variable

R = residual

---

[1]The RCC5V computer program adjusts for unequal cell sizes.

Stepwise regression is a powerful variation of multiple regression which provides a means of choosing independent variables which will provide the best possible prediction of the dependent variable with the fewest independent variables. The method recursively constructs a prediction equation for one independent variable at a time. The first step is to choose the single variable which is the best predictor. The second independent variable to be added to the regression equation is that which provides the best prediction in conjunction with the first variable. This procedure continues by adding variables step-by-step until the desired number of independent variables have been added or until no other variables will make a significant contribution to the prediction equation. At each step the optimum variable is selected from those variables which remain.

In order for an independent variable to be added to the regression equation its regression coefficient (b, unnormalized) must be significant as measured by the F statistic and its tolerance level must be high enough to demonstrate that a new dimension is being added to the prediction equation. Consequently, stepwise regression never brings a variable into the equation if the tolerance is below a specified minimum.

Planned comparisons were done on all statistically significant interaction results which were predicted by the hypotheses. Post hoc analyses were done on all significant results which were obtained but not hypothesized. The Neuman-Kuels test was used for the comparison of means involved in these interactions.

CHAPTER III

RESULTS AND INTERPRETATION

The results are presented here under three main headings: results from between trials scores, results based upon the delayed criterion tests and results based upon data concerned with the establishing of mathemagenic behaviors. A brief interpretation of the results follow the presentation of each set with a more general discussion of the relationships between the three sets of results being reserved for Chapter IV. A correlation matrix of all independent and dependent variables may be found in Appendix F.

Between Trials Results

The mean score and standard deviation for each of the four experimental groups is presented in Table 2. The control group received no between trials questions and thus has no between trials score. Each mean is based upon 10 trials for 18 subjects. Inasmuch as a correct response on any trial would earn 2 points, the highest possible score for any subject was 20 points with the lowest being 0 points.

An analysis of variance was carried out with question type (remembering and inferring), question position (pre and post), location of information in paragraph (either first half or second half of the paragraph) and experimenter as the main factors. This was a 2x2x2x6 repeated measure. The results are summarized in Table 3.

In this analysis the main effect of question type was significant ($F = 11.01$, $df = 1/48$, $p < .005$). Those subjects who received remembering questions did significantly better than those receiving inferring questions. Though this result was not surprising, it was of relatively little impor-

Table 2

Means and Standard Deviations of Between Trials Total Score

Group	Mean	SD
Remembering Pre-Question	10.17	5.37
Remembering Post-Question	11.56	4.87
Inferring Pre-Question	4.78	4.17
Inferring Post-Question	9.78	4.40

Table 3

Analysis of Variance of Between Trials Total Score

Source	df	MS	F	p
Question Type (T)	1	112.01	11.01	<.005
Question Position (P)	1	88.67	8.72	<.005
Experimenter (E)	5	5.52	<1	
Location of Information in Section (I)	1	1.56	1	
T x P	1	27.56	2.71	n.s.
T x E	5	4.29	<1	
P x E	5	2.76	<1	
T x I	1	.01	<1	
P x I	1	39.06	14.92	<.001
E x I	5	2.11	<1	
T x P x E	5	42.75	4.20	<.005
T x P x I	1	2.51	1.16	n.s.
T x E x I	5	2.59	<1	
P x E x I	5	4.25	1.62	n.s.
S(T x P x E)	48	10.17		
T x P x E x I	5	2.19	<1	
SI(T x P x E)	48	2.62		

tance here as the effect of question type was only of interest in this study as it interacted with criterion type. The result does suggest, however, that these inferring questions as a group, were probably more difficult than the set of remembering questions.

The analysis also showed the main effect of question position to be significant ($F = 8.72$, $df = 1/48$, $p < .005$) with those subjects receiving post-questions getting significantly higher scores ($\bar{X} = 10.67$) than those subjects who received pre-questions ($\bar{X} = 7.47$). This finding was in the opposite direction of what was predicted.

The hypothesis that children would do better on those items where the information necessary to answer the question was close, as opposed to distant from the question, was not supported as there was no significant main effect due to the location of information in the paragraph ($F = < 1$, $df = 1/48$). The interaction between question position and the location of answer in the section, however, was highly significant ($F = 14.92$, $df = 1/48$, $p < .001$; see Figure 2 for the graph of this interaction). Results of the planned comparison test showed that subjects did better on items where they received a post-question and the information needed to answer the question was in the last half of the section than on items where they received a pre-question and the information was in the first half of the section ($p < .005$). On items where the information needed to answer the question was presented in the first half of the section, subjects who received post-questions performed better than those who received pre-questions ($p < .05$). On items where the information needed to answer the question was presented in the last half of the section, subjects who received a

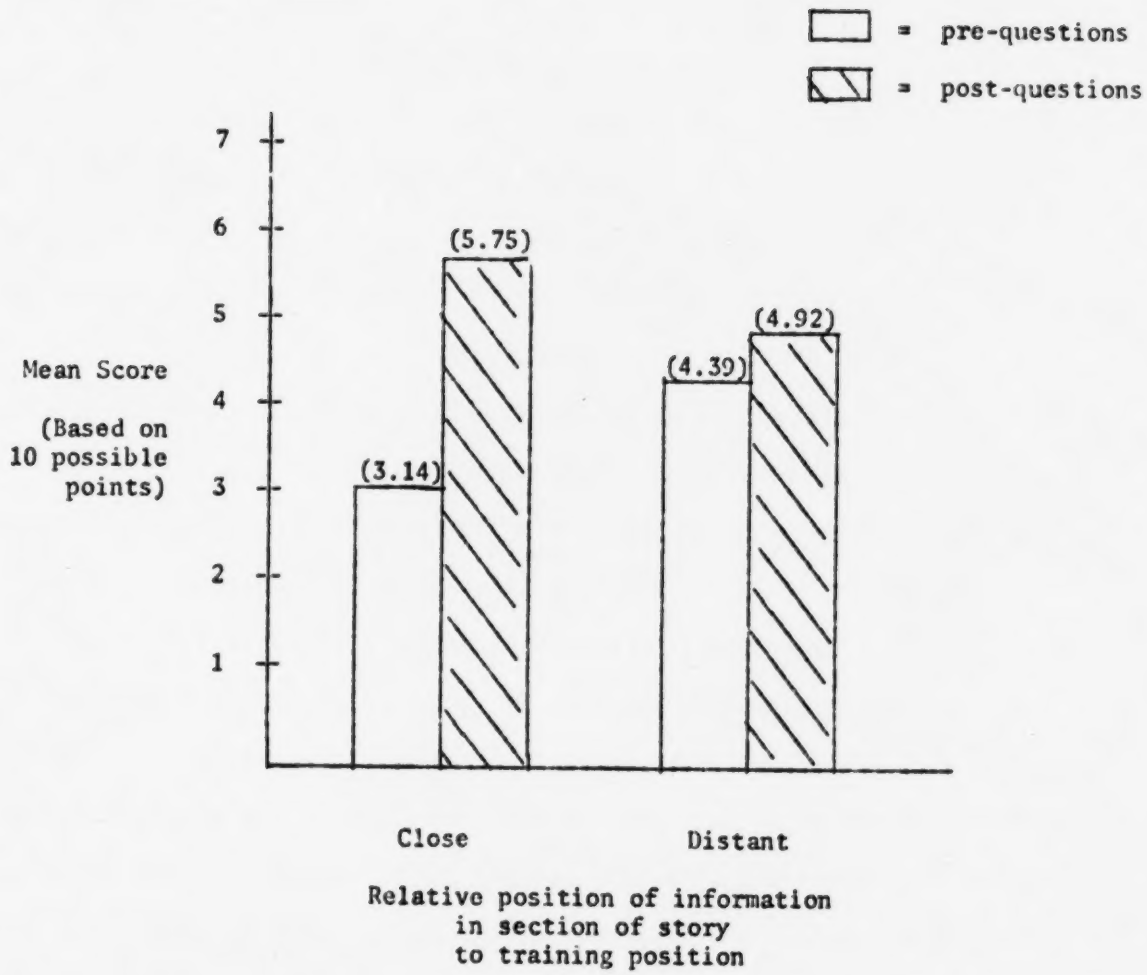

Figure 1 Graph of means of the interaction between location of information in sections of the story and position of question.

post-question out performed those subjects who had received pre-questions ($p < .05$).

Results of the analysis of variance also showed the interaction of question type, question position and experimenter to be significant ($F = 4.20$, $df = 5/48$, $p < .005$). See Figure 3 for the graph of this interaction. The Neuman-Keuls test was used in order to establish which means were significantly different from each other. Results of this comparison showed experimenter 5 (E_5), when asking post-remembering questions, to be able to get higher achievement from the children than experimenters 1, 2, and 6 when they asked pre-inferring questions ($p < .01$). The subjects of E_5 who received post-remembering questions also did better than the subjects of E_3, E_4, E_5 who received pre-inferring questions ($p < .05$), better than the subjects of E_5 and E_6 when they received post-inferring questions ($p < .05$) and better than E_5 and E_6 when their subjects received pre-remembering questions ($p < .05$). Children who received post-inferring questions from E_2 did better than children who received pre-inferring questions from E_1, E_2, or E_6 ($p < .05$). The children who received post-remembering questions from E_6 did better than the children of E_2 who received pre-inferring questions ($p < .05$) and subjects who received post-inferring questions from E_1 did better than those children who received pre-inferring questions from E_2 ($p < .05$).

These differences between experimenters may be accounted for in several ways. A first possibility is that chance alone may have accounted

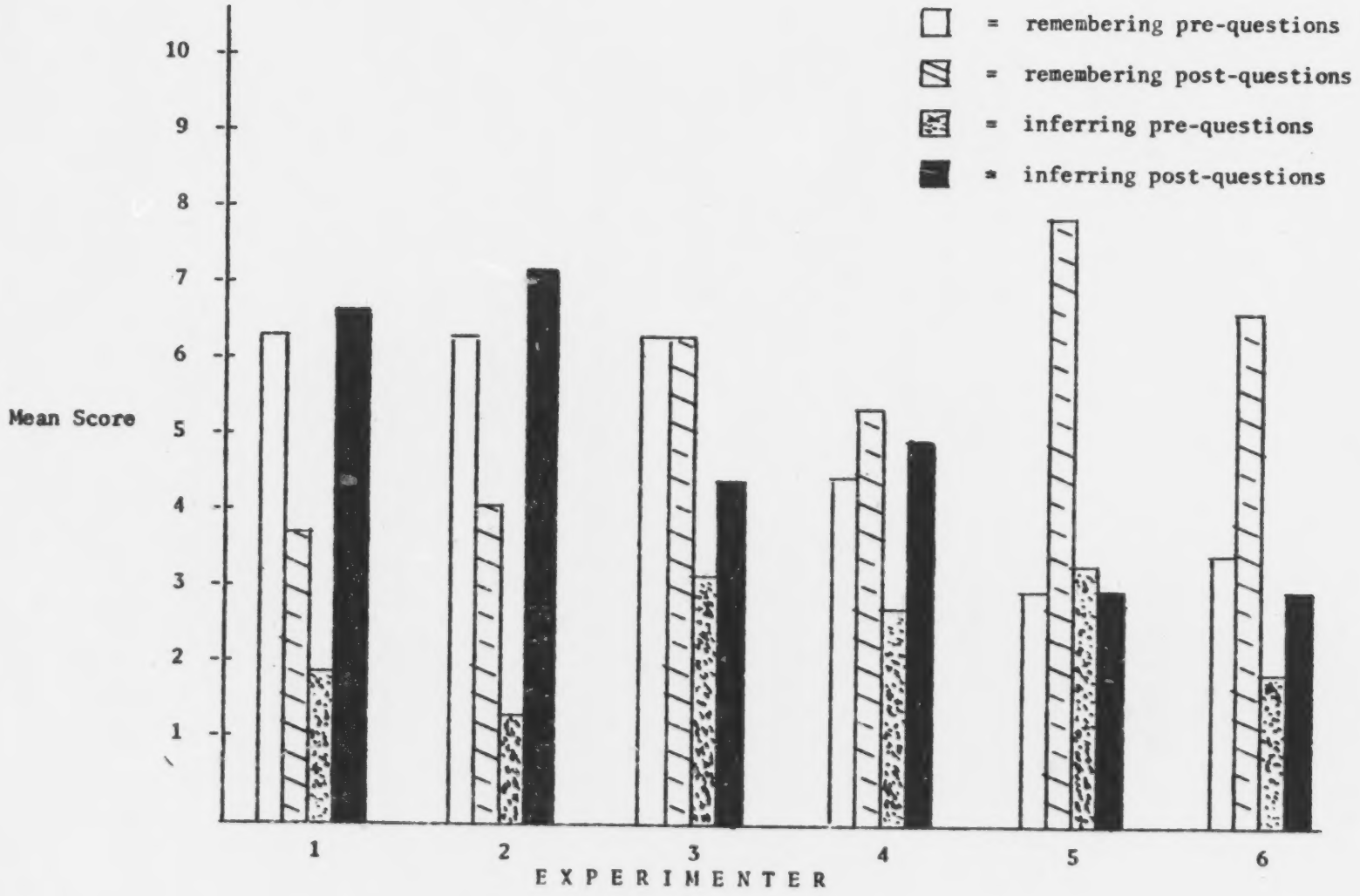

Figure 2. Graph of the mean scores involved in the interaction of question type x question position x experimenter (T x P x E)

for the observed differences. A second is that although a t-test of means showed the experimental groups not to be significantly different from each other with regard to mental age, it is possible (though unlikely) that they were not randomly assigned to treatment groups and, in fact, differed on some variable which was not measured. The third explanation which could be offered is that there were in fact real differences in the behavior of experimenters which accounted for the significance of the 3-way interaction. While this might be true, of the 15 means which were found to be significantly different, differences among 9 of these were likely due to the factor of criterion type. The remaining 6 may indeed be due to experimenter effects, but it would be risky at best to attempt to hypothesize the exact nature of these experimenter behaviors. The lack of any other significant interactions involving experimenters, the lack of a main effect involving experimenters, and the lack of any clear trend or influence of experimenters in the 3-way interaction would make this sheer speculation.

A stepwise multiple regression analysis was done on each of three dependent variables: (1) subjects' scores on items where the question and information needed to answer the questions were in close proximity to each other, (2) subjects' scores on tiems where the question and information necessary to answer the question were distant from each other, and (3) the sum of the two preceding scores or the total between-trials score.

The following independent variables could be included in the regression equation depending upon the significance of the variable (F statistic) and its tolerance level: question type, question position, mental

age, chronological age, sex, race of child, race of school and socio-economic status of the school which the child was attending. The results of these regression analyses are summarized on Tables 4, 5, and 6.

In looking at Table 4 (prediction of close proximity scores) it can be seen that question position, mental age and question type account for over 41% of the variance with question position and question type together accounting for 3 times as much variance as mental age. Though both question type and question position showed significant main effects in the analysis of variance, it seems that question position is the more powerful variable of the two, at least in its relationship to scores on close proximity items. Sex, race of the child and socio-economic status reached statistical significance but do not appear to add appreciably to the precision of the regression equation. Predominant race of the school attended by the child and chronological age were not included in the equation.

In predicting scores on distant proximity items (Table 5) mental age and question type continue to account for approximately 10% of the variance each, but question position contributes only slightly more than 1.5% of the total variance. This is a drop from almost 21% of the variance in predicting close proximity scores. A total of 31% of the variance was accounted for in predicting scores on those items where the question and information needed to answer the question were separated by at least half of the section.

In predicting the total between-trials score (Table 6), 39.39% of the variance was accounted for with all variables entered being included in the regression equation. Question type, mental age and question position

Table 4

Stepwise Multiple Regression Analysis on

Close Proximity Scores

Independent Variable	R Square	B	Standard Error B	F (at last step)
Question Position	.2074	2.8402	.5468	
Mental Age	.3069	0.0725	.0206	
Question Type	.4188	-1.9555	.5570	
Sex	.4206	0.3103	.5768	
Race of Child	.4229	0.2788	.5740	
SES of School	.4232	-0.0953	.5590	7.95*
(Constant)		-4.7457		

*df = 6/65, $p < .001$

Variables not included:

Race of School

Chronological Age

Table 5

Stepwise Multiple Regression Analysis on

Distant Proximity Scores

Independent Variable	R Square	B	Standard Error B	F (at last step)
Mental Age	.1113	.0839	.0273	
Question Type	.2156	-2.0766	.6384	
SES of School	.2556	-1.2735	.6748	
Race of School	.2805	1.2629	.8824	
Question Position	.2983	.7198	.6193	
Chronological Age	.3081	.0314	.0339	
Race of Child	.3115	.4816	.8578	4.13*
(Constant)		3.6627		

*df = 7/64, $p < .001$

Variables not included:

Sex of Child

Table 6

Stepwise Multiple Regression Analysis on

Total Between Trials Score

Independent Variable	R Square	B	Standard Error B	F (at last step)
Question Type	.1161	-4.0402	1.0804	
Mental Age	.2479	.1634	.0461	
Question Position	.3640	3.5872	1.0516	
SES of School	.3789	-1.2385	1.1411	
Race of School	.3876	1.1340	1.4932	
Chronological Age	.3930	-.0439	.0575	
Sex of Child	.3938	.3033	1.1081	
Race of Child	.3939	-.1696	1.4602	5.12*
(Constant)		-.1117		

*df = 8/63, $p < .001$

continued to account for the major share of the accountable variance (36.40%) with each variable now contributing approximately equal amounts.

These data on short term retention and achievement suggest question position to be an important variable in influencing the learning of educable mentally retarded children. As a main effect subjects who received post-questions did consistently better than those who received pre-questions. This effect, though significant, was in a direction opposite to that of the prediction. It was hypothesized that the pre-question condition condition would cue the subject to listen for the appropriate information (i.e., the answer) and thus enhance his retention of the critical information. This effect, it was thought, would be superior to the effect of the expectation held by others who were receiving post-questions, that they would be asked a question about the section of the story they had just heard.

The significant interaction between question position and location of the critical information in the section of the story seems to offer data in support of an explanation for this finding. See Figure 2 (p. 44) for the graph of this interaction. A comparison of the scores on the post-question close-proximity items with the pre-question close-proximity items provides the clearest example of the interaction. Subjects seemed to do better when both the question and information were in close temporal proximity to the time when the response was called for. The group which had both the question and critical information furthest removed from the call for the response did the worst. A comparison of the means involved in this interaction shows that in no case did a pre-question group perform better

than a post-question distant group and the pre-question close group offers further support for the importance of the question position variable. Though the appropriate interaction of question position and location of information provides the most power in influencing retention, location of critical information relative to the request for the response to the question is not a powerful variable when taken alone. This is further supported by the lack of a significant main effect of location of information in the analysis of variance (Table 3).

A further comparison of the interaction means suggests that the information in the story may be more easily remembered than the questions which preceded the individual sections. The significant difference between the post-question distant items and the pre-question close proximity items in favor of the post-question distant items suggests that even when the information necessary to answer the question is furthest removed from the request for the response, subjects do better than when the question precedes the section and is quickly followed by the response.

The pre-question, close proximity situation seems to offer the greatest opportunity for the interference of new information to influence forgetting. Thus the information which follows the question, but before the answer serves to interfere with the retention of the question. And then the information which follows the critical information (i.e., the answer) again serves to interfere with the retention of this critical information. In the post-question conditions this interference between information and question presentation does not have a chance to operate. The explanation of the findings offered by this interference hypothesis

is further supported by the difference (though not statistically significant) in favor of the post-question close-proximity items over the post-question distant proximity items.

The application of the short-term memory deficit (Ellis, 1963, 1970) hypothesis may also be used to explain these results as subjects (1) did consistently better in the post-question condition, (2) showed better retention where the question was a post-question and the information necessary to answer the question was in the second half of the section as opposed to the pre-question group with the information in the first half of the section. In this last group the greatest amount of time exists between the time of the presentation of the information and the request for the response; thus, creating the greatest opportunity for forgetting. The former group of course, creates the least opportunity for forgetting.

In order to test which of these hypotheses offers the best explanation for the observed phenomenon, further research would need to be done. By controlling the nature of the activity between the presentation of the critical information and the call for the response (i.e., by continuing to present information or just allowing time to lapse) one would be able to identify the relative effects of forgetting through interference, and forgetting through the passage of time alone.

In summary, it seems clear that question position is an important variable in influencing the short term retention of educable mentally retarded children. And, moreover, that the best results on an immediate retention test may be obtained under conditions where there is close temporal contiguity between the question and the presentation of the information required to provide the correct answer.

Delayed Critertion Test Results and Interpretation

Results to be reported in this section are concerned with subjects' performance on the delayed (after 15-minute rest period) criterion test. There were 20 items on this criterion test. Ten of these items were "relevant" items -- the questions which were asked by the experimenter during the listening activity, and 10 were "incidental" items -- questions relating to the material covered in the story but not asked during the listening activity. Thus the criterion test was composed of 2 subtests.

Though the control groups received all 20 items as the criterion test, it must be pointed out that in fact, they have no "relevant" items. This is a result of their having received no training questions during the listening activity. The same situation exists for those groups who received one type of training question and another type of question on the criterion test. Because the questions they received on the criterion test were of a different type than those they heard during the listening activity, they really have no "relevant" criterion test questions either. These clarifications in mind, it must be remembered that the term relevant is used to describe a certain subset of 10 items on the criterion test and not any treatment conditions. The same applies for the term incidental which refers to the remaining set of 10 items. These 10 items were never heard by any of the subjects but nevertheless pertain to information presented in the 10 sections of the story.

A 2x2x5x2 analysis of variance was used to compare the relevant delayed criterion test scores without the control group. The main factors involved in this analysis were question type (remembering and inferring), question position (pre and post), experimenter (6 levels) and criterion

type (remembering and inferring). The means and standard deviations for each of the ten experimental groups may be found in Table 7. Each mean was derived from 18 subjects and based upon 10 questions.

The results of this analysis, summarized in Table 8, showed a significant main effect of criterion type ($F = 10.71$, $df = 1/24$, $p < .005$). Closer examination of the means indicated that subjects did significantly better on the remembering criterion test than on the inferring criterion test. This finding was consistent with results of the between trials analysis where a significant difference was found with respect to question type.

The results of the analysis also showed a significant interaction between question type and criterion type ($F = 9.70$, $df = 1/24$, $p < .005$). See Figure 3 for a graph of this interaction. Results of a planned comparison test of means showed that subjects who received remembering training questions and a remembering criterion test performed better than those subjects who received inferring training questions and the remembering criterion test ($p < .005$); and that subjects who received inferring training questions and an inferring criterion test did better than those subjects who received remembering training questions and the inferring criterion test ($p < .05$). These findings were in the predicted direction and confirm hypotheses 10 and 11. Other means were also found to be significantly different from each other: subjects who received remembering training questions and a remembering criterion test did better than the group which received remembering training questions and inferring criterion test ($p < .005$) and the group which received inferring training questions and inferring criterion test ($p < .01$), and subjects who received inferring training questions and a remembering criterion test did better than those subjects who

Table 7

Means and Standard Deviations of Relevant

Delayed Criterion Test Scores

(Based on a possible total of 20 points)

Group	Mean	SD
Criterion Test 1 (Remembering)		
Remembering Pre-Questions	13.55	3.71
Remembering Post-Questions	11.67	5.12
Inferring Pre-Questions	7.88	4.40
Inferring Post-Questions	9.22	3.23
Control (No Questions)	10.22	4.42
Criterion Test 2 (Inferring)		
Remembering Pre-Questions	5.44	2.92
Remembering Post-Questions	6.11	1.83
Inferring Pre-Questions	8.78	3.53
Inferring Post-Questions	8.66	4.69
Control (No Questions)	8.11	5.46

Table 8

Analysis of Variance of Relevant Delayed Criterion Test Scores without Control Group[1]

Source	df	MS	F	p
Question Type (T)	1	7.56	<1	
Question Position (P)	1	1.17	<1	
Experimenter (E)	5	19.88	1.13	n.s.
Criterion Type (C)	1	189.06	10.71	<.005
T x P	1	8.51	<1	
T x E	5	19.41	1.10	n.s.
T x C	1	171.17	9.70	<.005
P x E	5	4.02	<1	
P x C	1	.06	<1	
E x C	5	7.75	<1	
T x P x E	5	11.26	<1	
T x P x C	1	27.56	1.56	n.s.
T x E x C	5	8.62	<1	
P x E x C	5	18.45	1.05	n.s.
T x P x E x C	5	16.25	<1	
Error	24	17.65		
TOTAL	71			

[1]These data were analyzed in a 2x2x2 ANOVA as well, eliminating the experimenter factor. This was done in order to allow the RCC5V program to analyze the same data using equal n's. The results of this analysis showed the criterion type main effect to be significant at the $p < .005$ level as well as the TxC interaction ($p < .005$). All other effects were nonsignificant.

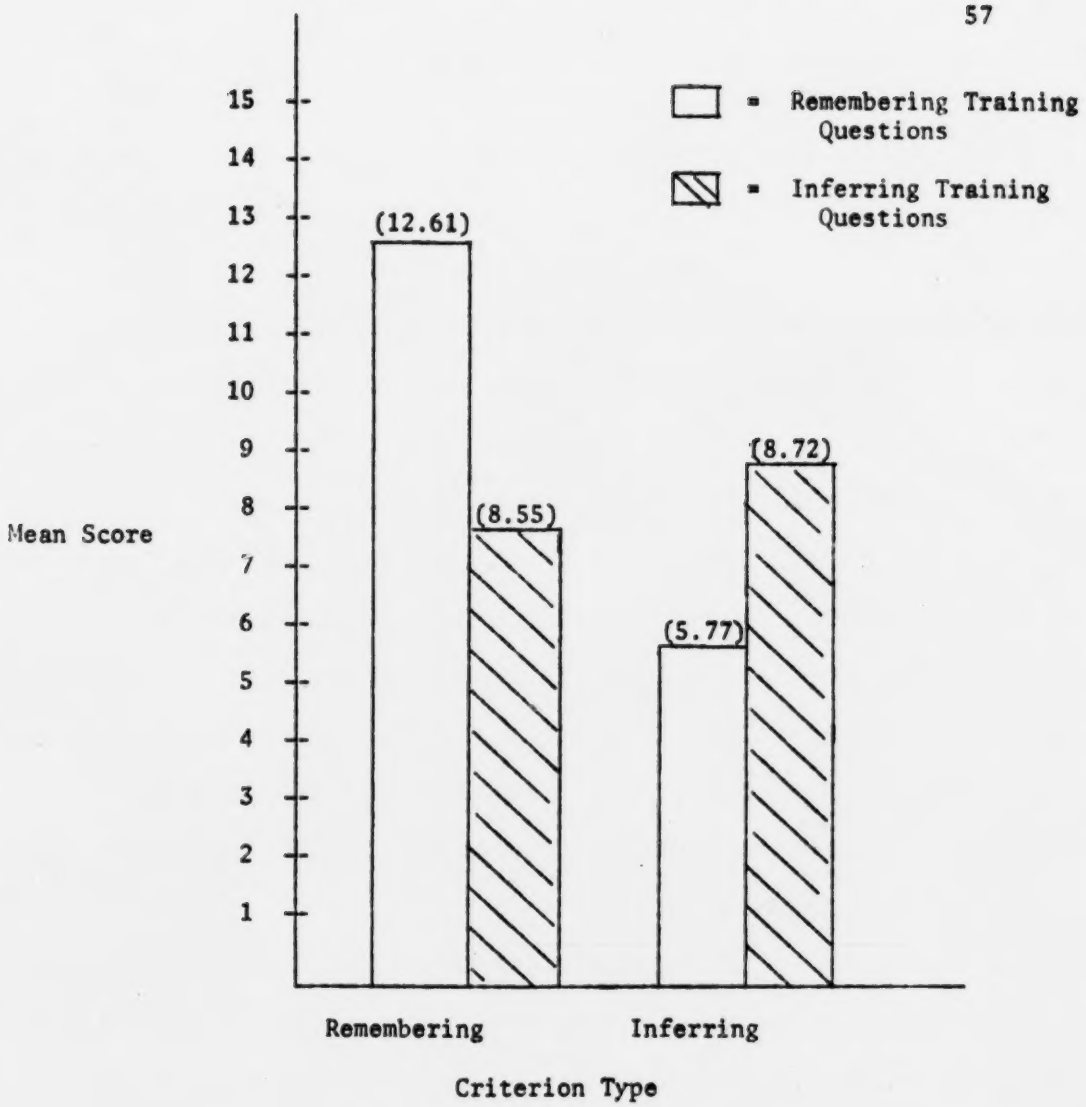

Figure 3. Graph of the mean scores involved in the interaction of training question type and criterion type (T x C) Relevant Score

received remembering training questions and an inferring criterion test ($p < .05$). These last three comparisons, however, though significant, may be accounted for by differences in apparent difficulty in the criterion test and thus are not of particular interest here.

Hypothesis 6 predicted that children who received pre-questions would do better on the relevant criterion test than children who received post-questions. The lack of a significant main effect of question position requires the rejection of this hypothesis.

A 6x6 analysis of variance was carried out in order to compare the relevant criterion test results with the control groups included. The main factors were 6 levels of experimenter and 6 levels of group. With the exception of the control groups, who of course received no training questions, only those subjects who received the same type of criterion question as training question were included in this analysis of 36 experimental subjects and all 18 of the control subjects.

The results of this analysis (as summarized in Table 9) indicated that there were no significant differences between groups, experimenters or in the interaction of these two variables. Thus Hypothesis 4, which predicted that children who received questions would do better on the relevant criterion test which asked the same questions as were in the training, than children who were in the control group and did not receive training questions, was not supported by the data. Yet previously presented data suggest that there is indeed an interaction between question type and criterion type. A careful inspection of the means (Table 7) offers a possible explanation of these two findings.

Table 9

Analysis of Variance of Relevant Delayed Criterion Test Scores with Control Group

Source	df	MS	F	p
Groups (G)	5	42.31	1.73	n.s.
Experimenter (E)	5	29.53	1.21	n.s.
G x E	25	15.63	<1	
Error	18	24.44		
TOTAL	53			

While the experimental groups which received the same type training question as criterion question did better than the control subjects, those subjects who received training questions and criterion questions which were not of the same type seem to have done worse than the respective control groups. Thus, while it can not be said that questioning in general produces better relevant retention after the 15 minute delay, it does seem clear that the best retention is obtained under conditions where the questions which are asked are of the same type as the criterion test questions. Thus, if questions are to be asked during instruction, it seems clear that attempts must be made to maintain consistency between the level of questions asked during the training period and those to be included on a criterion test.

Table 10 contains the means and standard deviations of all incidental delayed criterion test scores. Each of these scores is derived from 9 subjects and based upon 10 items.

A 2x5x2 analysis of variance was carried out on the relevant and incidental delayed criterion subtest scores with the control groups included. The main factors were criterion type (remembering and inferring), groups (5 levels) and subtests (relevant and incidental). A summary of these results may be found in Table 11. As before, the criterion type main effect was found to be significant ($F = 7.50$, $df = 1/16$, $p < .05$) with children who received the remembering criterion test doing better than those who received the inferring criterion test. Significant interactions were also found between criterion type and groups ($F = 4.80$, $df = 4/64$, $p < .01$) and groups and the subtests ($F = 3.42$, $df = 4/64$, $p < .05$).

With respect to the interaction of criterion type and groups, a test of planned comparisons showed the following means to be significantly dif-

Table 10

Means and Standard Deviations of Incidental Delayed Criterion Test Scores

(Based on possible total of 20 points)

Group	Mean	SD
Criterion Test 1 (Remembering)		
Remembering Pre-Questions	5.44	2.88
Remembering Post-Questions	6.33	3.46
Inferring Pre-Questions	6.33	2.69
Inferring Post-Questions	7.00	4.56
Control (No Questions)	6.11	3.41
Criterion Test 2 (Inferring)		
Remembering Pre-Questions	3.56	3.28
Remembering Post-Questions	6.11	5.37
Inferring Pre-Questions	7.67	3.77
Inferring Post-Questions	6.89	4.96
Control (No Questions)	4.67	2.40

Table 11

Analysis of Variance of Relevant and Incidental Delayed Criterion Subtest Scores with Control Groups

Source	df	MS	F	p
Criterion Type (C)	1	133.47	7.50	$<.05$
Groups (G)	4	5.91	<1	
Subtests (RI)	1	29.61	<1	
Error $_{S(RI)}$	16	50.19		
C x G	4	58.75	4.80	$<.01$
C x RI	1	2.45	<1	
G x RI	4	45.13	3.42	$<.05$
Error $_{CS(RI)}$	16	17.78		
Error $_{GS(RI)}$	64	13.21		
C x G x RI	4	13.89	1.14	n.s.
Error $_{CGS(RI)}$	64	12.24		

ferent from each other: on the remembering criterion test children who received remembering pre-questions did better than those children who received inferring pre-questions ($p < .05$); on the inferring criterion test, the children who received inferring pre-questions did better than those children who received remembering pre-questions ($p < .05$) and children who received inferring post-questions did significantly better than those children who received remembering pre-questions ($p < .05$). See Figure 4 for a graph of this interaction. Significant differences were also found between criterion types in the pre-question condition, with those receiving the remembering criterion test doing better than those who took the inferring criterion test ($p < .005$). And again, the pre-question group which received the remembering criterion test did better than the remembering post-question group that received the inferring criterion test ($p < .05$). These last two interactions were thought to be significant as a result of the unequal difficulty of criterion test items, and thus, not of great interest here. No significant differences were found in comparing treatment groups with the control group on the same criterion test type for either the remembering or inferring criterion tests, or between control groups.

In looking at the interaction of group by subtest (relevant and incidental), the planned comparision test of means showed several differences to be significant (See Figure 5 for graph of this interaction). Those subjects who received inferring post-questions did significantly better on the incidental criterion test than did those who received the same type of training question on the relevant criterion test ($p < .005$) and those who received inferring pre-questions on the relevant test ($p < .05$). In comparing control

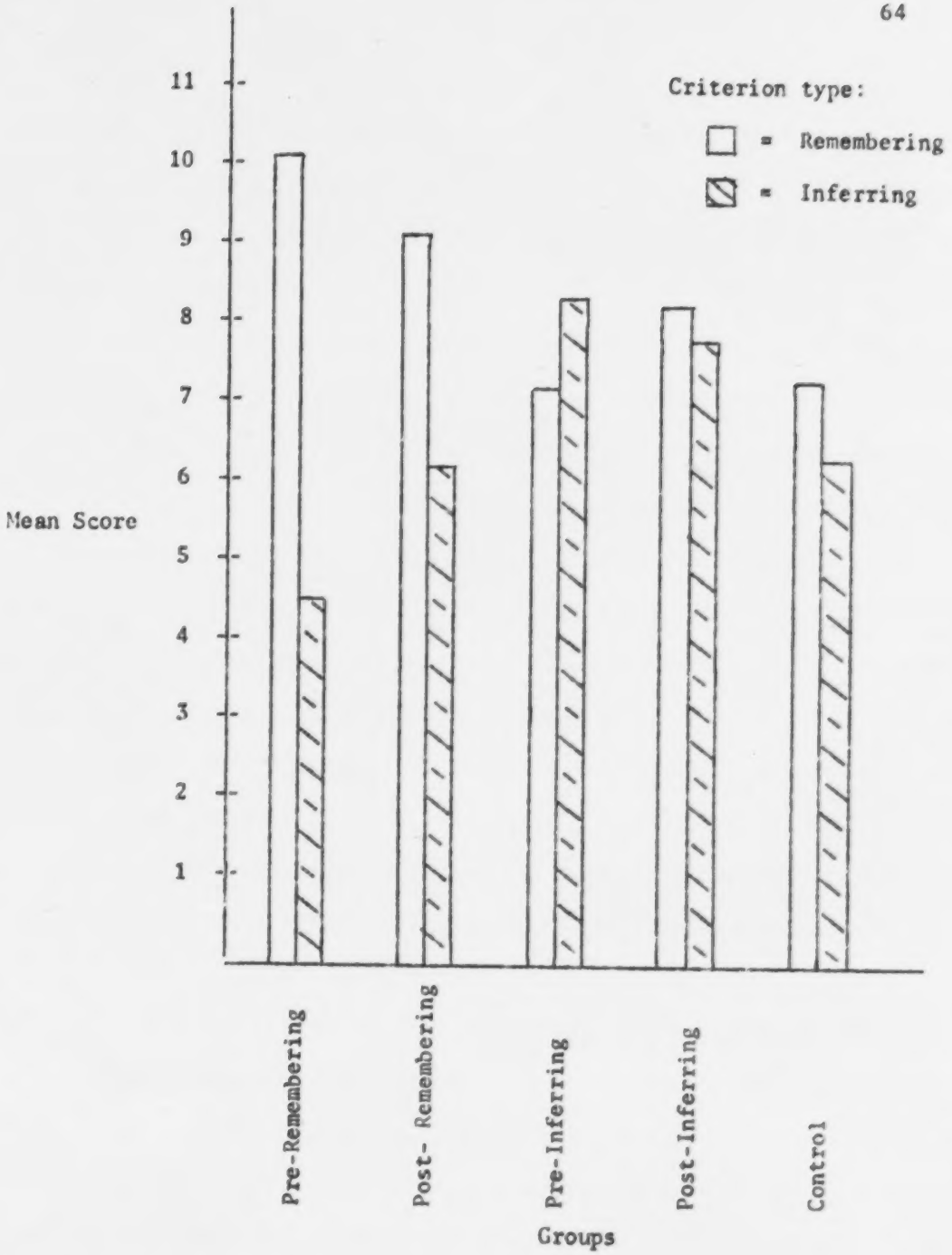

Figure 4. Graph of the interaction between criterion type and groups.

(C x G)

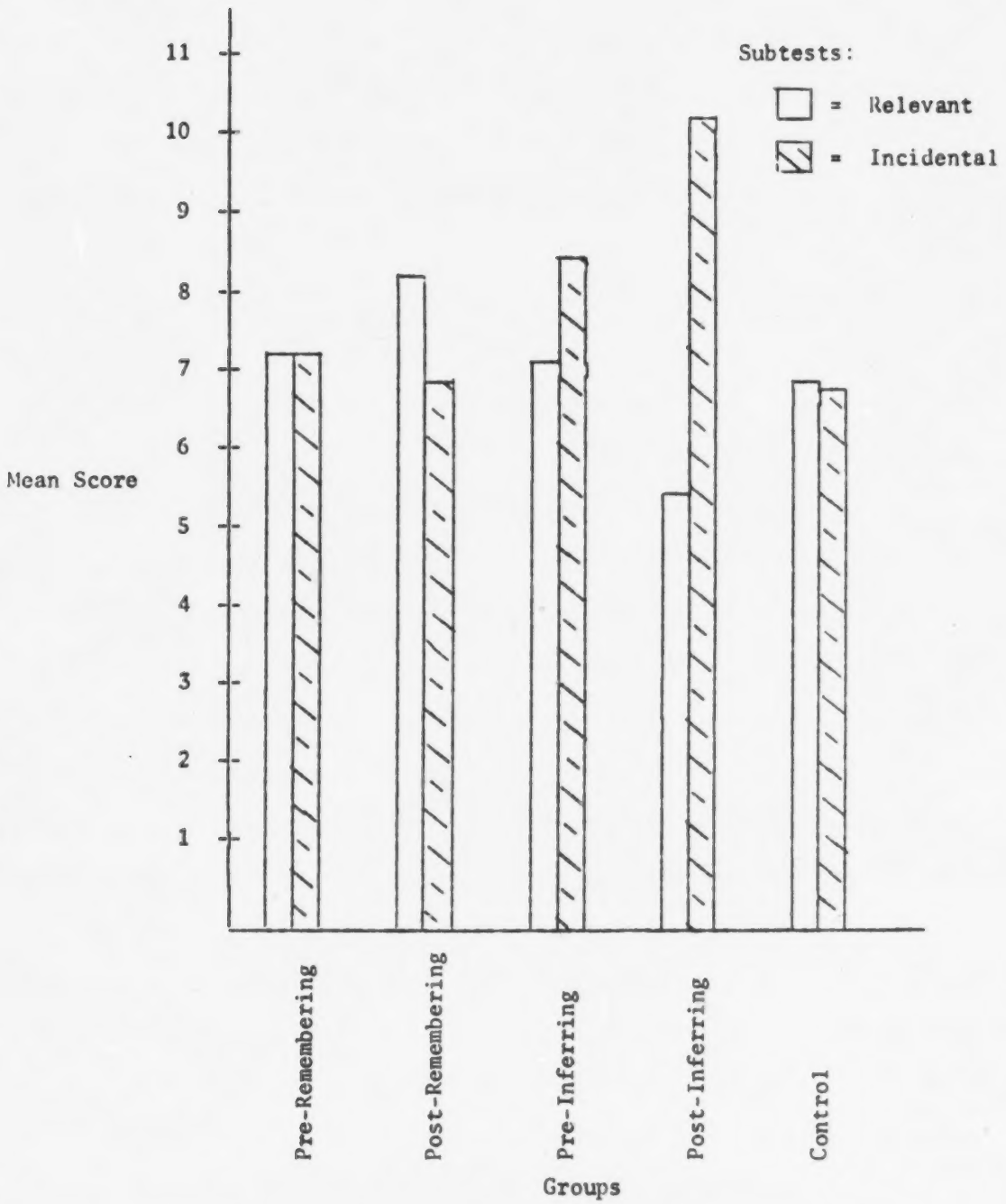

Figure 5. Graph of interaction between groups and subtests on delayed criterion test (G x RI)

groups on the relevant and incidental criterion tests, no significant differences were found. This finding suggests the relevant and incidental subtests to be of equal difficulty. The test of means also showed the post-inferring group to do significantly better on the incidental subtest than either the post-remembering group ($p < .005$), the incidental control group ($p < .005$) or the pre-remembering group ($p < .01$) did on the incidental criterion subtest. In comparing relevant subtest scores, the post-remembering group was found to have done significantly better than the post-inferring group ($p < .05$).

A close look at the graph of this interaction (Figure 5) shows practically a mirror image difference between the relevant and incidental scores. With the exception of the pre-question remembering group and the control group, the scores are just the opposite of each other with respect to treatment groups. The most interesting results are the differences between the post-inferring groups on the incidental subtest as compared with all the others. The difference between this group on the incidental and relevant subtests, for example, suggests that by asking post-inferring questions one cannot hope to increase both incidental and relevant learning and that, in fact, in asking post-inferring questions one significantly increases the incidental learning while tending to suppress relevant learning (difference between this group and control on relevant learning in favor of control group though not statistically significant).

The significantly better performance of the post-inferring group on the incidental subtest as compared with the performance of the control, pre-remembering and post-remembering groups on the incidental subtests suggests

the asking of post-inferring questions to be an effective technique for increasing incidental learning.

The asking of post-remembering questions also seems to be a better way of producing relevant learning than the use of post-inferring questions. The fact that the relevant means of these two groups differ significantly from each other but not from the relevant control mean suggests a relatively small facilitating effect for the post-remembering questions and a relatively small suppressing effect for the post-inferring questions on relevant learning.

Table 12 contains the summary data from the 2x2x6x2 analysis of variance data using total criterion test score as the dependent variable. The main factors in this analysis were question type (2 levels), question position (2 levels), experimenter (6 levels) and criterion type (2 levels). A significant main effect of criterion type was found ($F = 4.26$, $df = 1/24$, $p < .05$) with children continuing to do better on the remembering than on the inferring criterion test. A significant interaction between training question type and criterion type was also indicated ($F = 4.54$, $df = 1/24$, $p < .05$). See Figure 6 for the graph of this interaction. The planned comparison test of means showed the remembering training question - remembering criterion type group to have done significantly better than the remembering training question - inferring criterion type group ($p < .005$). This again was thought to have been due primarily to differences in degree of difficulty between criterion types. The mean of the remembering training question - inferring criterion test group was also found to be significantly lower than the means of both the inferring training question - inferring criterion test group ($p < .05$) and the inferring training question - remembering criterion

Table 12

Analysis of Variance of Total Delayed Criterion

Test Scores without Control Group

Source	df	MS	F	p
Question Type (T)	1	6.67	< 1	
Question Position (P)	1	10.56	< 1	
Experimenter (E)	5	43.97	< 1	
Criterion Type (C)	1	232.56	4.26	< .05
T x P	1	2.51	< 1	
T x E	5	73.87	1.35	n.s.
T x C	1	248.06	4.54	< .05
P x E	5	19.73	< 1	
P x C	1	.56	< 1	
E x C	5	38.96	< 1	
T x P x E	5	62.64	1.15	n.s.
T x P x C	1	98.34	1.80	n.s.
T x E x C	5	30.43	< 1	
P x E x C	5	58.63	1.07	n.s.
T x P x E x C	5	33.91	.62	
Error	24	54.65		
TOTAL	71			

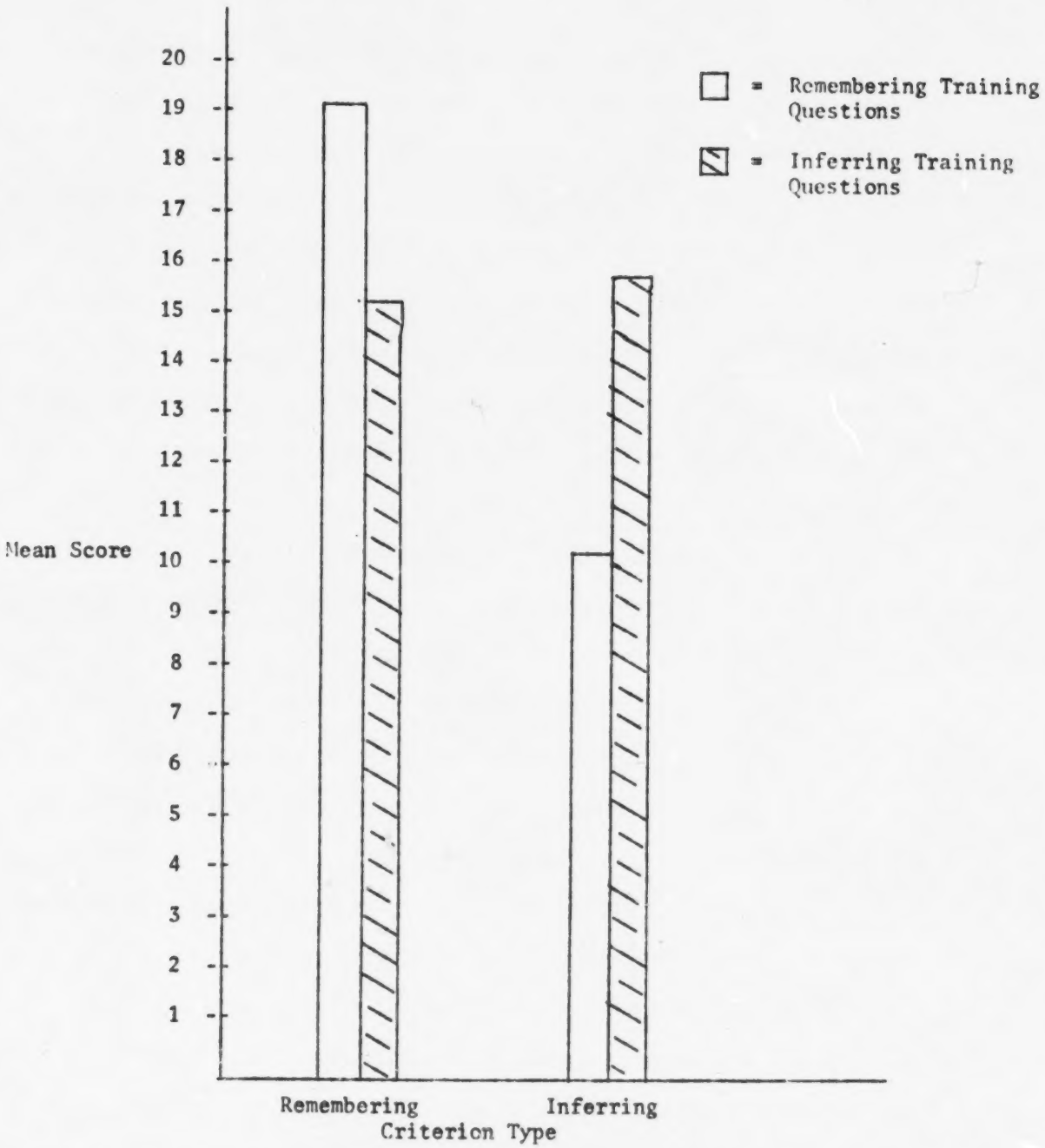

Figure 6. Graph of the mean scores involved in the interaction of training question type and criterion type on total criterion score

type group ($p < .05$), thus offering partial confirmation of hypothesis 10 -- that the total score criterion test results would be highest for children receiving the same type of criterion test as training question.

The results of analyses in this section necessitated the rejection of two hypotheses. A planned comparison test of the control group incidental means and training question groups' incidental means indicated no significant differences and thus required the rejection of hypothesis 5 -- that children who received questions would do better on the incidental criterion test than children who were in the control group and did not receive training questions.

The lack of a main effect difference for subtests (RI) in Table 11 required the rejection of hypothesis 9 -- that children would do better on the relevant subtest than on the incidental subtest.

Tables 13, 14, and 15 contain the results of the stepwise multiple regression analysis on the dependent variables of relevant delayed criterion test scores, incidental delayed criterion test scores and total delayed criterion test scores respectively.

The results of the regression analysis on the relevant criterion test scores (Table 13) showed that a total of 41.54% of the variance was accounted for with all of the 9 independent variables entered. Mental age accounted for the greatest amount of variance for a single variable (15.19%) with criterion type accounting for approximately 13% of the variance. Question position accounted for the least amount of variance of all the variables entered (.06%).

In predicting the incidental criterion test scores (Table 14) the multiple regression analysis accounted for a total of 32.76% of the variance.

Table 13

Stepwise Multiple Regression Analysis on Relevant

Delayed Criterion Test Scores

Independent Variable	R Square	B	Standard Error B	F (at last step)
Mental Age	.1519	0.1552	.0381	
Criterion Type	.2816	-3.4550	.8762	
Chronological Age	.3670	-0.1121	.0481	
SES of School	.4033	-1.6789	.9412	
Question Type	.4088	-0.6677	.8918	
Sex	.4116	0.5913	.9140	
Race	.4141	0.6974	1.2189	
Race of School	.4148	-0.3406	1.2369	
Question Position	.4154	0.2155	.8676	4.90*
(Constant)		16.4287		

*df = 9/62, $p < .001$

Table 14

Stepwise Multiple Regression Analysis on Incidental Delayed Criterion Test Scores

Independent Variable	R Square	B	Standard Error B	F (at last step)
Mental Age	.1710	.1414	.0365	
Sex	.2530	-1.8994	.8773	
Race of School	.2771	1.8737	1.1874	
Chronological Age	.2979	-0.0650	.0462	
Question Type	.3077	0.8714	.8561	
Race of Child	.3156	-1.1581	1.1701	
Criterion Type	.3222	-0.6424	.8412	
Question Position	.3267	0.5531	.8329	
SES of School	.3276	-0.2509	.9035	3.36*
(Constant)		2.5011		

*df = 9/62, p < .005

Table 15

Stepwise Multiple Regression Analysis on Total Delayed Criterion Test Scores

Independent Variable	R Square	B	Standard Error B	F (at last step)
Mental Age	.2118	.2966	.0649	
Criterion Type	.2687	-4.0974	1.4938	
Chronological Age	.3440	-0.1771	.0820	
Sex	.3550	-1.3081	1.5581	
SES of School	.3671	-1.9298	1.6046	
Race of School	.3736	1.5331	2.1087	
Question Position	.3764	.7686	1.4791	
Race of Child	.3768	-0.4608	2.0781	
Question Type	.3770	0.2036	1.5203	4.16*
(Constant)		18.9299		

*df = 9/62, p < .001

Mental age continued as the single most powerful variable (17.10%) with sex contributing the second most (8.20%). Question type and position, though having statistically significant contributions added little to the precision of the regression equation. The same observation holds for the influence of criterion type.

The results of the regression analysis on the dependent variable of total score (Table 15) shows a total of 37.70% of the variance being accounted for. Mental age contributes over 21% to this total with criterion type being the second most important variable (5.79%). Again question type and question position, though making statistically significant contributions to the total variance accounted for, contribute little in terms of increasing the precision of the equation.

The most striking observation in comparing these regression analyses with those of the between trials scores is the difference in the power of the question position variable in the short term task relative to the delayed criterion task. It appears as if the more remote the criterion task from the learning task, the less important the position variable in influencing learning outcomes.

Results and Interpretation with Respect to Mathemagenic Behaviors

Two different methods were used in an attempt to test the hypothesis that questions may facilitate the development of mathemagenic behaviors and result in the establishing of learning sets or, more specifically, in this study, listening skills. It was predicted that if questions do have this facilitative effect then two things would happen. First, the subjects' performance on the between trials items would improve as the listening

task proceeded. And secondly, that the post-question conditions would have a more powerful influence on the incidental delayed criterion test scores than the pre-question conditions, when the criterion test was of the same type as the training questions.

The mean incidental score and standard deviation for each of the 8 experimental groups and 2 control groups is presented in Table 10. Each mean was derived from 9 subjects and based upon 10 questions. These were the 10 items in the 20-item criterion test which the subject had never heard before. They were not training questions though they did pertain to information covered in the story. Inasmuch as a correct response on any trial would earn 2 points, the highest possible score for any subject was 20 points with the lowest being 0 points.

A 2x2x6x2 analysis of variance was used to analyze these incidental delayed criterion test scores. The factors were question type (remembering and inferring), question position (pre and post), experimenter (6 levels), and criterion type (remembering and inferring). The results of this analysis are summarized in Table 16.

This analysis showed no main effects or interactions to be significant at the $p < .05$ level of significance. The main effect of question position did not show post-questions to produce better incidental score performance than pre-questions.

Nor was the prediction that the interaction of question type, question position, and criterion type would be significant found to be in the predicted direction. Thus, it must be concluded from those results that the training questions were not significant in influencing the development of mathemagenic

Table 16

Analysis of Variance of Incidental Delayed Criterion Test Score without Control Group

Source	df	MS	F	p*
Question Type (T)	1	28.44	1.77	n.s.
Question Position (P)	1	4.69	<1	
Experimenter (E)	5	7.14	<1	
Criterion Type (C)	1	2.25	<1	
T x P	1	1.78	<1	
T x E	5	23.06	1.43	n.s.
T x C	1	7.11	<1	
P x E	5	19.84	1.23	n.s.
P x C	1	.25	<1	
E x C	5	13.40	<1	
T x P x E	5	32.56	2.03	n.s.
T x P x C	1	21.78	1.35	n.s.
T x E x C	5	12.56	<1	
P x E x C	5	17.37	1.08	n.s.
T x P x E x C	5	9.06	<1	
Error	24	16.08		
TOTAL	71			

*p < .05

behaviors or listening skills. These results require the rejection of Hypothesis 15.

The results of the analysis of variance in Table 17 show the experimental groups not to be significantly different from the control groups with respect to the dependent variable of incidental delayed criterion test score. Thus, it appears that after the 15 minute delay between listening to the story and the criterion test, the experimental groups which received training questions were at no great advantage in answering the delayed incidental criterion test questions than the control groups which received no training questions at all.

A serial analysis of between trials scores was done in an attempt to identify possible learning to learn curves. It was reasoned that if subjects were learning how to listen throughout the listening task, that their performance would imporve across trials. In addition, the serial analysis by trials could indicate if there was any depreciation in performance -- possibly as a function of the length of the task -- if the subjects' performance was poorer at the end of the listening task. While an analysis of variance was done on these data, and reported in Table 18, it must be noted that the results of this analysis should be viewed with caution for it was not established that all items (trials) were of equal difficulty. Indeed, there is evidence to suggest that all the items were not of equal difficulty (e.g., main effect difference of question type in Table 3) and all of the effects significant in Table 18 may be accounted for by differences in item difficulty. As a result, it appears to be of greater value to look at the relationships between question position given the same question type. These relationships are graphed in Figures 7, 8, and 9.

Table 17

Analysis of Variance of Incidental Delayed Criterion Test Scores with Control Group

Source	df	MS	F	p*
Criterion Type (C)	1	10.00	<1	
Groups (G)	4	14.64	1.02	n.s.
C x G	4	11.31	<1	
Error	80	14.41		
TOTAL	89			

*p < .05

Table 18

Serial Analysis of Variance of

Between Trials Scores

Source	df	MS	F	p
Between	71	2.80		
Groups (G)	3	15.22	6.76	< .001
Error (G)	68	2.30		
Within	648	.70		
Trials (T)	9	2.80	4.65	< .001
G x T	27	2.23	3.69	< .001
Error (T)	612	.60		
TOTAL	719	.91		

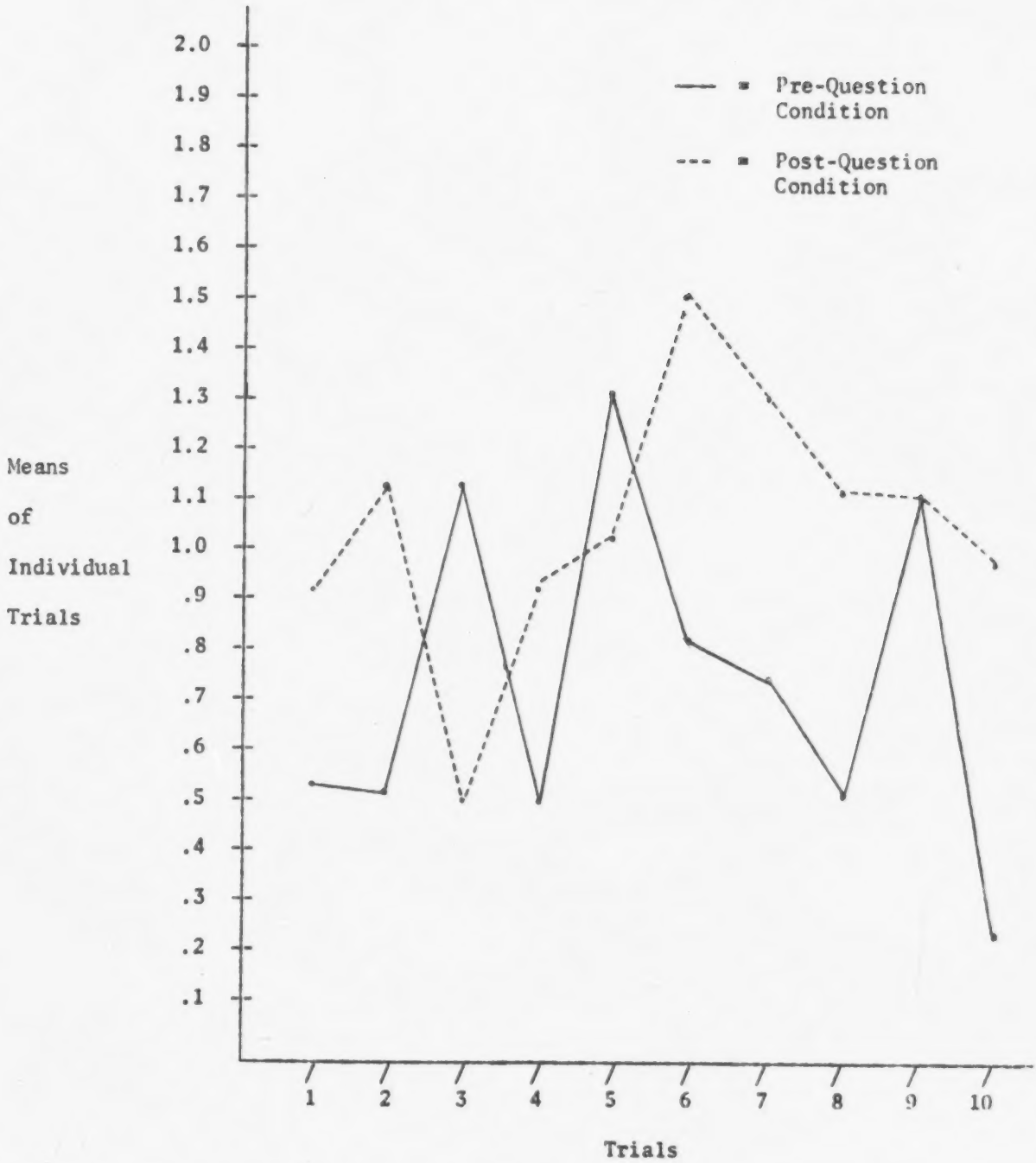

Figure 7. Graph of means of serial analysis by trials for pre- and post-question conditions.

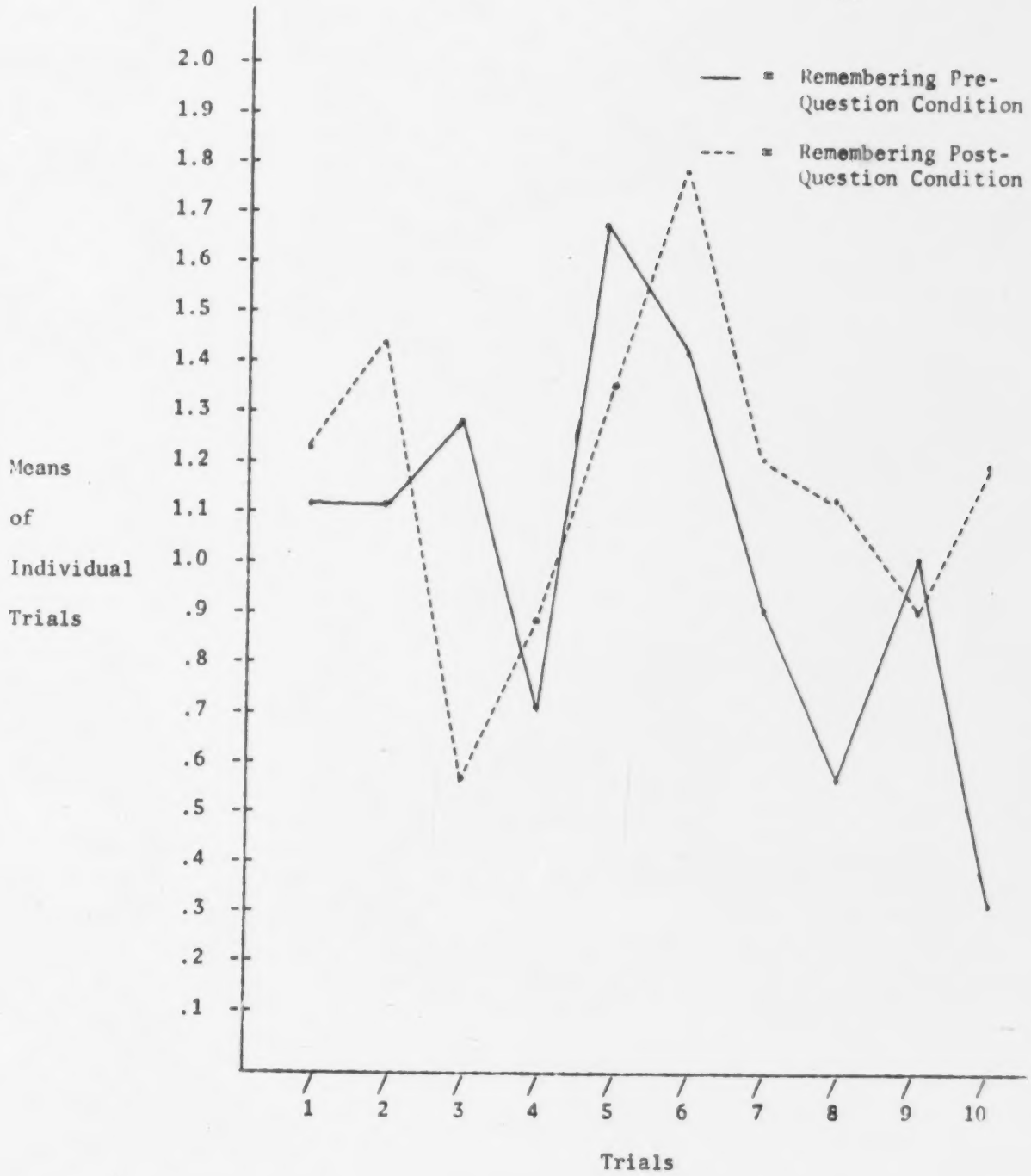

Figure 8. Graph of means of serial analysis by trials for remembering pre- and post-question conditions.

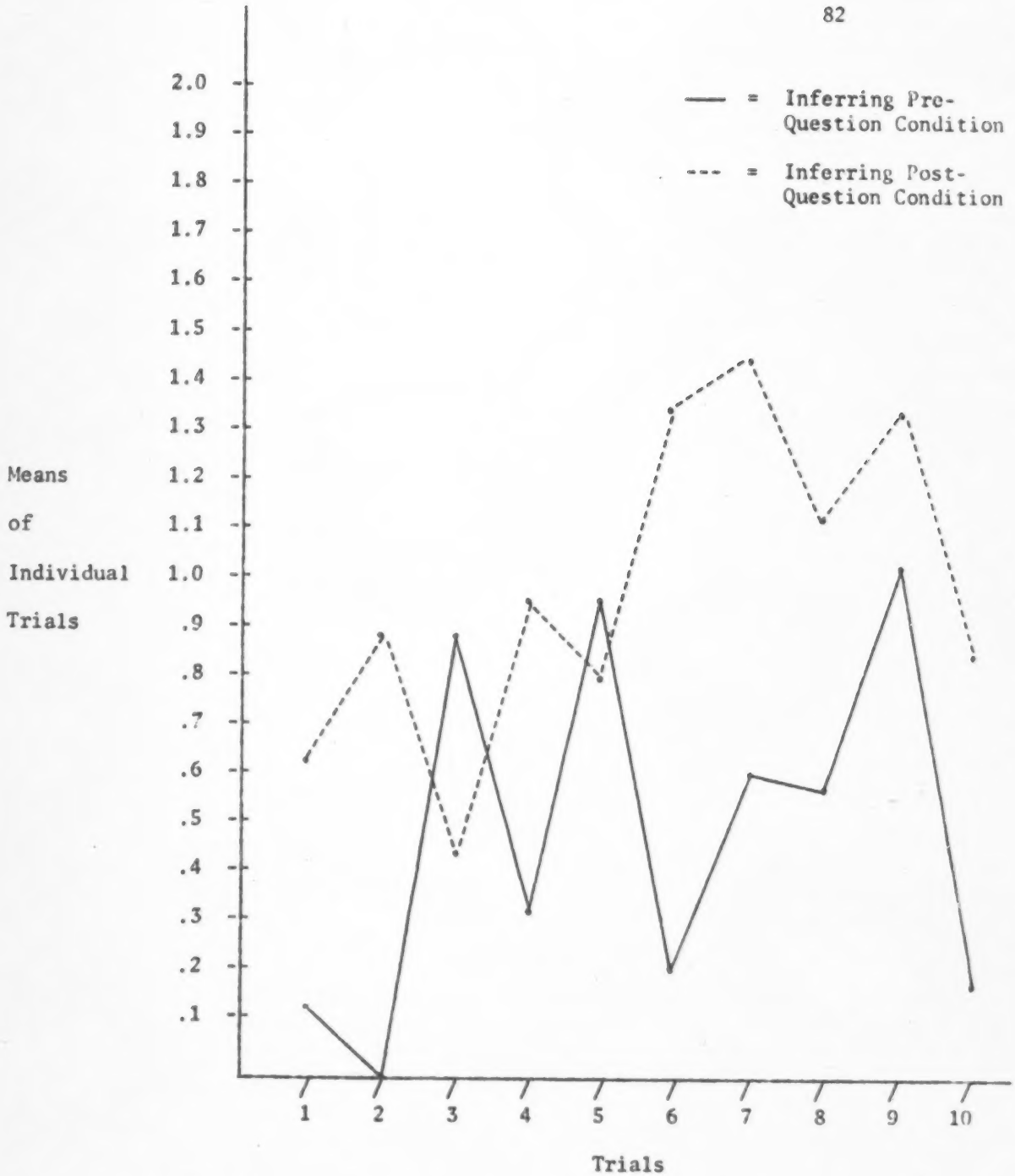

Figure 9. Graph of means of serial analysis by trials for inferring pre- and post-question conditions.

The graph of the serial analysis by trials for pre- and post-question conditions (Figure 7) reveals the possible beginning of the development of a listening skill in the post-question condition as predicted. Success on trials six through ten appears to be greater than on trials one through five for the post-question groups. A closer examination of this phenomenon (see Figures 8 and 9) shows that this improvement may be largely accounted for by improved scores in the post-inferring condition as opposed to the post-remembering condition. Given that the inferring questions seem to have been more difficult than the remembering questions and that the post-questions produced better between trials total scores, it may have been that the inferring questions when paired with the post- condition were easy enough to allow the subjects to learn how to deal with them and, thus, to improve their scores over trials. The pre-inferring condition may have been just too difficult to allow improvement of scores over trials to improve by mere exposure alone. A better test of this learning to learn hypothesis would be available in situations where items were of equal difficulty across trials. Moreover, increases in the length of the listening task with an increased number of trials may give the learning skill a greater chance to develop. Increasing the number of trials and decreasing the length of each trial may also be a productive way to test the hypothesis.

CHAPTER IV

DISCUSSION

This chapter will present a theoretical discussion and summary of the results derived from the present study and the implications these results have for the teaching of the mentally handicapped.

Discussion of Results

Hypothesis 1 of this study predicted that children receiving pre-questions would perform better than children who received post-questions on the between-trials total score (short-term retention). This was based on the rationale that questions could be effectively used to high-light information and increase short-term retention by cueing the children to the appropriate information to be remembered. The data did not confirm this hypothesis, but, in fact, showed just the opposite to be the case.

The prediction was made in Hypothesis 2 that children would score higher on trials where the information necessary to answer the questions was in "close proximity" as opposed to "distant proximity" from the question. The analysis of variance (Table 3) did not indicate a main effect due to location of information in the section and thus, Hypothesis 2 must be rejected as well.

Hypothesis 3 predicted that children would perform better on trials where they received pre-questions and the information was close, i.e., in the first half of the section, as compared with trials where a post-question was given and the information was close, i.e., in the last half of the section. Results of the analysis of variance showed an interaction between question position and location of information in the section -- but in a direction opposite to the prediction. Results indicated that subjects did

better on items where they received a post-question and the information needed to answer the question was in the last half of the section than on items where they recieved a pre-question and the information was in the first half of the section.

These findings were thought to be explained best by interference theory. Forgetting, according to interference theory, occurs because competing responses learned before the acquisition of criterion responses (proactive interference) or in the retention interval (retroactive interference) somehow interfere with the "habit" of the criterion responses. This interference produces a decrement in criterion behavior called forgetting.

In the present study the better post-question condition performance of subjects, as compared with subjects in the pre-question condition may be accounted for by the retroactive interference of the great amount of information presented after the question. Thus by the time the experimenter called for the response to the question in the pre-question condition, the subjects may well have forgotten the question. In the post-question condition, remembering the question was no problem. This explanation makes the assumption that the meaningfulness and interest of the story made it more easily remembered than the relatively sterile and weak presentation of the question.

Interference theory may also be used in explaining the interaction results where the subjects performed better on the items which were post-questions and where the information needed to answer the question was in the last half of the section, than on any other combination of the two

variables. In looking at Figure 1 (p. 42), the graph of this interaction, there seems to be a clear relationship between subjects' performance and the opportunity for interference to occur. Thus, subjects do best on items where there is the least opportunity for interference; the worst where there is the greatest opportunity for interference; and their performance is at a point in between these two groups when the opportunity for interference is at a point half way between these two extremes (i.e., on the distant proximity items where the question and critical information are separated by half of the listening passage).

Most of the research used to support interference theory is based upon the verbatim recall and recognition of unrelated words and nonsense syllables. However, a recent review of literature in this area (Cunningham, 1972) concludes that those studies reviewed support the assertion that the verbatim recall and recognition of prose follows the same laws of verbal interference established for the verbatim recall and recognition of unrelated words and nonsense syllables. It is further contended that, though the research findings are still equivocal, the weight of the evidence seems to indicate that verbatim retention and meaning retention are comparable and follow the same laws.

The application of short-term memory theory (Ellis, 1963, 1970) also offers a possible explanation for these findings. However, the application of short-term and long-term memory research conclusions would involve the application of findings generated from studies which have used primarily artificial learning tasks of much shorter duration than the listening task involved in this present study. The definitional problems of the

short-term and long-term memory concepts also imposes restrictions and limitations on the application of such findings to the present study.

However, casting these reservations aside for the moment, one may speculate that the subjects performed best in the post-question close proximity condition because it placed the fewest demands on them for stimulus organization, rehearsal strategies and secondary memory. Research reported in the first chapter has shown mentally retarded children to be deficient with respect to each of these variables when compared to their normal IQ peers.

In order to test which of these orientations, if either, offers the best explanation for the findings, additional data must be collected. By controlling the intervening activity between the presentation of the question and the information needed to answer the question, or in the post-question situation, by controlling the activity between the presentation of the information first, followed by the question, one may be able to determine the relative effects of forgetting due to interference of competing information as compared with the effects of time alone.

The multiple regression analyses provide further evidence for the importance of better understanding the interaction of question position and location of information. The regression analysis on the total between trials score (Table 6, p. 49) shows mental age and question type to account for approximately the same amount of variance -- 12%. However, this analysis does not consider the location of critical information in the section of the story. If this variable is considered the relative importance of question position and mental age change when predicting the distant proximity scores,

i.e., where the critical information and question are separated by half of the section, mental age accounts for over 11% of the variance with question position accounting for scarcely more than 1.5%. However, when the question and information are in close proximity to each other, the effect of question position is vastly increased -- accounting now for over 20% of the variance with mental age continuing to contribute approximately 11%. The general conclusion to be drawn from these analyses, then, seems to be that as the proximity between the question and the critical information increases, the more important becomes the time and activity occuring between the time of learning and the opportunity to give the response. (The reader should keep in mind that the post-question information close proximity group performed the best on the short-term retention test.)

In considering the delayed-criterion dependent measures, the following hypotheses were not supported by the data: Hypothesis 6 - that children who received questions would score higher on the relevant criterion test which asked the same questions as were in the training, than children who were in the control group and did not receive training questions; Hypothesis 7 - that children who received pre-questions would perform better on the relevant criterion test than children who received post-questions; and Hypothesis 9 - that children would perform better on the relevant subtest than on the incidental subtest.

These findings suggest that any differences which may have been obtained on the short-term retention measure were lost by the time the delayed criterion tests were given 15 minutes later. Several explanations may be offered for these results.

First, the lack of any differences between the groups which received training questions and the control groups on the relevant and incidental criterion tests may be accounted for by the nature of the instructions given to the control group. Those children who participated as control subjects were strongly encouraged to listen very carefully to the story and told that they would be asked questions about the story some time after it was over. While this would have provided the strongest test for the effect of questioning, it, combined with the obvious effects of participating in a study (Hawthorne effect), may have served to sufficiently boost the scores of the control group to make it appear as if the training questions had little long term effect. Efforts to eliminate this effect must be made in future attempts at creating control groups for studies of this kind. Moreover, it should not be concluded that the control group represented a reasonable facsimile to the "real classroom" and that questions therefore make no difference. The fact that each child received individual instructions to pay close attention to a story which was intended to have high interest value and was read by an enthusiastic professional story teller, is a clear departure from the everyday routine of the classroom.

A second possible explanation, and one which would include the lack of any signficant difference between the relevant and incidental criterion scores, is that the children may have been quite tired after the 15 minute rest activity plus the 15-20 minutes it took to get through the story and questions. This would have served to produce a poorer performance than might otherwise have been expected on a task of shorter duration.

A third possible explanation may be that the length of each section of the story reduced the effectiveness of the training questions and thus reduced the power of the treatment overall. Studies which shortened the length of the sections of the information to be learned would be likely to increase the number of between trials items the subject got correct. This would then increase the likelihood of his remembering the correct response to the relevant items through a straight-forward stimulus-response association.

The lack of a significant interaction effect between criterion type and question type in the analysis of variance of incidental delayed criterion test scores also forced the rejection of two additional hypotheses: Hypothesis 12 - that children who received remembering training questions would score higher on the remembering incidental criterion test than children who had received inferring training questions; and Hypothesis 13 - that children who received inferring training questions would perform better on the inferring incidental criterion test than children who had received remembering training questions.

The analysis of variance on the relevant delayed criterion test scores (Table 8) showed a significant interaction effect between question type and criterion type and thus provided confirmation of Hypothesis 10 - that children who received remembering training questions would score higher on the remembering relevant criterion test than children who had received inferring training questions and Hypothesis 11 - that children who received inferring training questions would perform better on the inferring relevant criterion test than children who had received remembering training

questions. The analysis of variance on the total criterion score (Table 12) produced a similar result and, thus, gave partial confirmation to Hypothesis 12 - that the total score criterion test results would be highest for children receiving the same type of criterion test as training question. (The relationship held only for the case where the inferring training question group had done better on the inferring criterion test than the remembering training question group on the same test; those who received remembering training questions did not perform significantly better than those who had received inferring training questions on the remembering criterion test.)

These results were taken as support for the general thesis that the activities a student engages in when confronted with instructional tasks are of critical importance in determining what he learns. The alternative view of course would be that the student is a passive receptacle whose learning and performance are directly determined by input variables. The reader interested in this debate is directed to Anderson's 1970 article on "Control of student mediating processes during verbal learning and instruction."

In general, then, it can be said that different types of questions may only be useful and necessary to the extent to which they facilitate the "desired" kind of learning. If one wanted children to perform well on a remembering criterion test, then inferring training questions would obviously be inappropriate. And, if one wanted a child to do well on an inferring criterion test, then remembering questions would be out of place as part of the training. It is not completely clear, however, whether this effect is a function of the facilitating effect of the appropriate type question relative to the criterion task, or due to the suppressing effect of the in-

appropriate type question. A close inspection of the means (Table 7) indicates that the inappropriate type question group performed slightly worse than the control group in each case. It should also be recalled that the relevant group did not perform significantly better than the control group on the delayed criterion test. Thus, the issue of the relative facilitating effects of the similar type training questions and criterion questions versus the possible suppressing effect of the dissimilar type training and criterion question remains clouded. Further research will be needed to unravel this important and perplexing issue.

Nevertheless, this significant interaction effect suggests what Rosenshine and Furst (1971) have called the "cognitive process" opportunity-to-learn phenomenon. This phenomenon refers to the important consideration of whether or not the level of the criterion instrument was relevant to, and of the same type as, the instruction. Overall, Rosenshine and Furst (1971) report, the correlations between measures of opportunity to learn and student achievement have been positive, significant, and consistent. This claim gained additional support with the results of a recent study (Watts and Anderson, 1971) which demonstrated that high school students did best on the criterion test when the questions they received were of the same type as those in the test.

The results of the analysis of variance of the incidental delayed criterion test scores (p. 76) indicated that there were no significant differences with respect to the question position - criterion type interaction, and, thus, led to the rejection of Hypothesis 15 - that children who received post-training questions would perform better than children who received

pre-training questions on the incidental criterion test where the items in the criterion test were of the same type as the training questions. This finding is not surprising given the clear lack of power and influence of the question position variable after the 15 minute delay. (Regression analysis (Table 14, p. 72) showed the question position variable to account for less than 1% of performance.) One must conclude from this evidence that the position of the training questions was not effective in facilitating the development of those mathemagenic behaviors necessary for the development of listening skills or learning sets. The fact that the development of listening skills was not supported by the delayed criterion test data, however, does not rule out the possibility that these learning sets were developed during the listening task and lacked the strength to persist over the period of the 15 minute interval.

The results of the serial analysis by trials were able to shed some light on this possibility. General confirmation of this possibility cannot be claimed. Though the fact that children who received pre-questions did not show improvement over trials (and thus led to the acceptance of Hypothesis 5), the results did not confirm the general improvement of children who received post-questions over trials, and thus led to the rejection of Hypothesis 4. The graph of this interaction, however, indicates the possible beginnings of the development of a learning set in the inferring post-question condition (see Figure 9, p. 82). The relatively better performance of the post-inferring questions as compared with the pre-inferring questions during the last half of the listening task and the better performance of the subjects on the last half of the post-inferring questions sug-

gests the beginnings of a learning to learn phenomenon. Research which increased the number of trials during a shorter period of time or increased the number of trials and lengthened the time of the task might shed further light on the development of these phenomena. Inasmuch as other research (e.g., Kaufman, 1971) suggests that the development of learning sets may be enhanced by overlearning, increased amount of training in the post-question condition may provide evidence for the development of listening skills through skillful questioning. One possibility for testing this hypothesis would be to have multiple training sessions over a period of days followed by the criterion test.

Implications for Teaching the Mentally Handicapped

Research reported in the review of the relevant literature has demonstrated that teachers ask many questions during classroom instructional activities. While the results of the present study cannot be used to support or reject the validity of this type of teacher behavior in total, they can be useful in suggesting that studies which record merely the frequency of questioning in the classroom seem to be missing essential ingredients in productive educational practice.

One of the results of this study suggests that the temporal proximity of a question in relation to the information to be learned is an essential relationship in the teaching of EMR children. The suggestion to teachers of the mentally handicapped, then, must be that their questioning practices be related to the information which is being presented in such a way that small units of information are quickly followed up by questions. If the time interval between the presentation of the information and the question

is too great or, if too much information is presented at once, the short-term effectiveness of the questioning will be severely reduced. Moreover, the findings suggest that questioning of this nature will be more effective when it follows the presentation of information as opposed to when it precedes it.

Teacher questioning effectiveness may take many forms. One of these is to facilitate the retention of information which a teacher might want students to remember. A second may be to draw attention to critical attributes in concept formation tasks. A third purpose for asking questions may be to give practice in different types of thinking.

Often it is assumed that retarded children are not capable of "higher levels" of information processing. The present study has not only demonstrated this not to be true, but has, in fact, shown that questions can be used to stimulate children to think in different ways and that questions can be used to produce different types of learning. This may be considered support for the general notion that the activities students engage in when confronted with instructional tasks are of crucial importance in determining what they will learn and further, that pupils are thus active agents in their own learning. Retarded children appear to be no different in this respect. However, in teaching the retarded, more attention must be given to the management of these student activities as they tend not to process information as effectively as normal children, or engage in appropriate self-control over their own study skill behavior. Questions can be useful educational tools to these ends.

Results of this study also suggest an important relationship between the type of question asked during instruction and the type of criterion

performance the teacher desires of the students. The type of question one asks seems to be important only to the extent that it is consistent with the type of learning desired. Desired student achievement would thus seem to be increased by teachers' being able to clearly state their instructional goals in terms of "cognitive process types" as well as in terms of content.

The findings also seem to have implications for the frequency of appropriate questioning of individual students; and especially of EMR students. Significant relationships have been demonstrated between the opportunity a child has to respond to questions and school achievement (Van Wagenen and Travers, 1963; Travers et al., 1964). Yet several studies show that children perceived as slower by their teachers tend to be slighted in classroom interactions by having fewer questions directed to them (e.g,, Brophy and Good, 1970; Lynch and Ames, 1971). If timely questioning is important for short-term retention (as this study has shown it to be), then it seems logical to conclude that children who are perceived as slow learners may not learn as much as they possibly could in classroom situations where they do not receive frequent questioning. This further supports the importance of small teacher-student ratios in the teaching of mentally retarded children as well as the need for frequent teacher-student verbal interaction.

The general conclusion of this study must be that questions can play an important role in the teaching of mentally handicapped children if the right kind of question is asked at the right time. The right type of question must be defined in terms of the desired learning outcomes and the right time is in close proximity and after the information to be learned has been presented.

CHAPTER V

SUMMARY

This study investigated the effects of question type and position on four types of learning among mentally handicapped children.

Sixty years of descriptive research had demonstrated that questions played an important role in the daily instructional activities of teachers. Further findings of these studies indicated that approximately two-thirds of the questions asked required direct recall of information presented in class. Studies of this type showed this percentage to be similar among various different levels of classrooms and subject areas. Additional evidence suggested similar percentages among special education classes.

While educators have been concerned with the type and frequency of questions asked in classrooms for a long time, and more recently with questions in special education classes, little research had been done on the effects of different types of questions.

Several studies indicated that educable mentally retarded (EMR) children were capable of productive thinking. Yet other research indicated that low-ability children and children for whom teachers had low expectations missed out on opportunities for intellectual stimulation from the teacher and that teachers frequently called on students they expected to give the right answer. Still, it remained to be demonstrated that questions were capable of being effective stimuli in providing opportunities for this kind of intellectual stimulation and productive thought in general and among EMR children in particular. This study represented one such effort.

Based on previous empirical studies concerned with question type, research on mathemagenic behavior (e.g., Rothkopf, 1966; Frase 1968a, 1969b), and relevant characteristics of educable mentally retarded children derived from theory and research (e.g., Ellis, 1963, 1970; Semmel, 1965; Spitz, 1966; Zeaman and House, 1963) fifteen different hypotheses were formulated. In general, these hypotheses predicted that pre-questions would produce better learning than post-questions; that questions would be most effective when the information needed to answer the question was close as opposed to distant; and that subjects would improve over trials. The data were all generated on between trials items and are referred to as the short-term retention test.

Hypotheses concerned with the delayed criterion test predicted that: subjects who received questions would perform better on the criterion test than control subjects who did not receive training questions; subjects who received the same type of criterion test question as training question would perform better than those subjects who received different types of training and criterion questions; subjects who received pre-questions would perform best on the relevant criterion test with subjects who received post-questions performing best on the incidental criterion test; and that subjects who received post-questions would show the development of a listening skill.

The subjects for this study were 90 intermediate EMR children. Each subject was asked to listen to a high interest-low vocabulary short mystery story of approximately 2400 words. This story was broken down into a total of ten sections.

Subjects were randomly selected from cooperating schools and classrooms and taken one at a time to the experimental room by the experimenter. After establishing rapport with the subject the experimenter indicated that they would be listening to a short story. The experimenter introduced the story with a general statement describing what the story was about. The subjects were tol that they would be asked some questions after they had finished the story.

During this listening activity the experimenter systematically introduced the questions. There were two types (remembering or inferring) in one of two positions (pre or post). In the pre-remembering question condition the experimenter asked a remembering question before the subject listened to each section. As was the case in all conditions the subjects were instructed to respond to the questions orally after listening to the section. The pre-inferring question group was asked one inferring question before each paragraph was heard. In the post- conditions the questions were asked one at a time, but after each paragraph was heard. There were a total of ten paragraphs and thus ten questions. A control group listened to the short story but without receiving any questions. All questions in both the treatment phases and on the cirterion test were of the free recall variety.

Th criterion test was administered 15 minutes after the experimental activity. This test was of two types and consisted of a total of 20 items. A subject could received either 20 remembering items or 20 inferring items. These 20 items were broken up into subtests of 10 items each. The relevant subtest consisted of items which were the same as those received during the listening activity. The remaining 10 items were of the same type and covered material included in the story though the subject had never received these items before.

The 15 minute rest activity between the end of the listening activity and the beginning of the criterion test consisted of playing with wooden blocks.

The data were analyzed using analysis of variance and stepwise multiple regression procedures. Results of the study indicated that important educational advantages may be derived from asking questions. On the short-term retention test, subjects who were asked questions in the closest temporal proximity to, and after the presentation of the critical information, were found to perform best. In addition to this, post-questions were found to produce better learning than pre-questions. While no statistically significant evidence indicated subject improvement over trials a trend in that direction was indicated for subjects who had received post-inferring training questions. The findings also indicated that approximately 12% of the variance on the short-term retention test (total score) was accounted for by the question position variable.

Results on the delayed criterion test indicated an interaction between training question type and criterion type with the best performance generally being obtained when the training questions and items on the criterion test were of the same type. No statistical evidence was found which supported the development of learning sets or listening skills. The results showed the question position variable to account for approximately 1% of the total variance on the delayed criterion test, thus indicating its diminishing effects over time.

Consistent differences were obtained with respect to the question type variable with subjects who received remembering items outperforming those who received inferring items.

Thus, support for all hypotheses was not indicated by the data analysis. Interference and memory theory and research were used to explain the findings.

Implications of the results for education were discussed with the general conclusion being that questions seem to offer great value for the teaching of the mentally handicapped when they are asked at the right time and with the right objective in mind. The right time was defined as being in close temporal proximity to the presentation of information and after the presentation of this information. Further, it was noted that the questions were of greatest value when they were of the same type as the type of question asked on the criterion test.

REFERENCES

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.

Belmont, J. B., and Butterfield, E. C. The relations of short-term memory to development and intelligence. In L.P. Lipsitt, and H. W. Reese (Eds.), Advances in Child Development and Behavior. Vol. 4. New York: Academic Press, 1969. pp. 29-82.

Berlyne, D. E. Structure and direction in thinking. New York: Wiley, 1965.

Brophy, J. E., and Good, T. L. Teachers' Communication of Differential Expectations for Children's Classroom Performance: Some Behavioral Data. Journal of Educational Psychology, 1970, 61, 365-374.

Bruner, J. S., Goodnow, J. J., and Austin, G. A. A Study of Thinking. New York: Wiley, 1956.

Bruner, J. S., Oliver, R. R., and Greenfield, P. M. Studies in Cognitive Growth. New York: Wiley, 1966.

Cawley, J. F., and Chase, D. V. Productive thinking in retarded and non-retarded children. British Journal of Educational Psychology, 1967, 37, 356-360.

Chall, J. S., and Feldman, S. C. A Study in Depth of First Grade Reading. New York: The City College of the City University of New York (U.S. Office of Education Cooperative Research Project No. 2728).

Connors, C. K., and Eisenberg, L. The Effect of Teacher Behavior on Verbal Intelligence in Operation Headstart Children. Baltimore: Johns Hopkins University School of Medicine (U.S. Office of Economic Opportunity Headstart Contract No. 510), 1966.

Cunningham, D. J. The retention of connected discourse: a review. Review of Educational Research, 1972, 42, 47-71.

Davis, O. L., and Tinsley, D. C. Cognitive objectives revealed by classroom questions asked by social studies teachers. Peabody Journal of Education, 1967, 45, 21-26.

Denny, M. R. Learning. In R. Heber and H. Stevens (Eds.), Review of Research in Mental Retardation. Chicago: University of Chicago Press, 1964. pp. 100-142.

Deutch, M. Unpublished paper delivered at Colloquium at Indiana University, 1966.

Ellis, N. R. Object-quality descrimination learning sets in mental derectives. Journal of Comparative and Physiological Psychology, 1958, 51, 79-81.

Ellis, N. R. The stimulus trace and behavioral inadequacy. Chapter in N. R. Ellis (Ed.), Handbook of Mental Deficiency. New York: McGraw-Hill, 1963, pp. 134-158.

Ellis, N. R. Evidence for two storage processes in short-term memory. Journal of Experimental Psychology, 1969, 80, 390-391.

Ellis, N. R. Memory processes in retardates and normals. Chapter in N. R. Ellis (Ed.), International Review of Research in Mental Retardation, Vol. 4. New York: Academic Press, 1970, pp. 1-32.

Ellis, N. R., and Hope, R. Memory processes and the serial position curve. Journal of Experimental Psychology, 1968, 77, 613-619.

Fine, M. J., Allen, C. A., and Medvene, A. M. Verbal interaction patterns in regular and special classrooms. Psychology in the Schools, 1968, 5, 265-271.

Floyd, W. D. An analysis of the oral questioning activity in selected Colorado primary classrooms. (Doctoral dissertation, Colorado State College) Ann Arbor, Michigan: University Microfilms, 1960. No. 60-6253.

Frase, L. T. Learning from prose material: length of passage, knowledge of results, and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.

Frase, L. T. Some unpredicted effects of different questions upon learning from connected discourse. Journal of Educational Psychology, 1968, 59, 197-201. (a)

Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249. (b)

Frase, L. T. Structural analysis of the knowledge that results from thinking about text. Journal of Educational Psychology Monograph, 1969, 60, Part 2, 1-16.

Furst, N. F. The multiple languages of the classroom. Paper presented at the meeting of the American Educational Research Association, February, 1967.

Gall, M. D. The Use of Questions in teaching. Review of Educational Research, 1970, 40, 707-721.

Gallagher, J. J. Expressive thought by gifted children in the classroom. Elementary English, 1965, 42, 559-568.

Goldstein, H., and Kass, C. Incidental learning of educable mentally retarded and gifted children. American Journal of Mental Deficiency, 1961, 66, 245-249.

Guszak, F. J. Teacher questioning and reading. The Reading Teacher, 1967, 21, 227-234.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), Psychology: A Study of a Science, Vol. 2. New York: McGraw-Hill, 1959, pp. 492-537.

Harris, A. J., and Serwer, B. Comparison of reading approaches in first grade teaching with disadvantaged children (The Craft Project). U.S. Office of Education Cooperative Research Project No. 2677. New York: City University of New York, 1960.

Harris, A. J., et al., A continuation of the CRAFT Project: Comparing Reading approaches with disadvantaged urban Negro children in Primary grades. U.S. Office of Education Project No. 5-0570-2-12-1. New York: Division of Teacher Education of the City University of New York, 1968.

Harter, S. Discrimination learning set in children as a function of IQ and MA. Journal of Experimental Child Psychology, 1965, 2, 31-43.

Harter, S. Mental age, IQ, and motivational factors in the discrimination learning set performance of normal and retarded children. Journal of Experimental Child Psychology, 1967, 5, 123-141.

Hebb, D. O. The Organization of Behavior. New York: Wiley, 1949.

House, B. J., and Zeaman, D. Miniature Experiments in the discrimination learning of retardates. In L. P. Lipsitt and C. C. Spiker (Eds.) Advances in Child Development and Behavior, Vol. 1, New York: Academic Press, 1963, pp. 313-374.

Hunkins, F. P. The influence of analysis and evaluation questions on achievement in sixth grade social studies. Paper presented at the annual meeting of the American Educational Research Association, 1967.

Hunkins, F. P. The effects of analysis and evaluation questions on various levels of achievement. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1968.

Kaufman, M. E. Long-term retention of a learning set in mentally retarded children. American Journal of Mental Deficiency, 1971, 75, 752-754.

Kaufman, M. E., and Peterson, W. M. Acquisition of a learning set by normal and mentally retarded children. Journal of Comparative and Physiological Psychology, 1958, 51, 619-621.

Kaufman, M. E., and Peterson, W. M. Acquisition of a conditional descrimination learning set by normal and mentally retarded children. American Journal of Mental Deficiency, 1965, 69, 865-870.

Kaufman, M. E., and Prehm, H. J. A review of research on learning sets and transfer of training in mental defectives. In N. R. Ellis (Ed.) International Review of Research in Mental Retardation. Vol. 2. New York: Academic Press, 1966. Pp. 123-149.

Kleinman, G. General Science Teacher's Questions, Pupil and Teacher Behaviors, and Pupils' Understanding of Science. Unpublished doctoral dissertation, University of Virginia, Charlottesville, Virginia, 1964.

Ladd, G. T., and Anderson, H. O. Determining the level of inquiry in teachers' questions. Journal of Research in Science Teaching, 1970, 7, 395-400.

Lynch, W. W., and Ames, C. Differences among special class teachers in individualization of instruction. Paper read at American Educational Research Association Meeting, New York, February, 1971. (a)

Lynch, W. W., and Ames, C. Individual cognitive demand schedule Center for Innovation in Teaching the Handicapped, Indiana University, Bloomington, Indiana. Technical Report No. 42, November 3, 1971.(b)

Lynch, W. W., and Ames, C. Personal communication. 1972.

Mandler, G. Organization and memory. In K. W. Spence and J. T. Spence (Eds.), The Psychology of Learning and Motivation. Vol. 1. New York: Academic Press, 1967, pp. 328-372.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Minskoff, E. H. Verbal interactions of teachers and mentally retarded pupils. Dissertation Abstracts, 1967, 27, 546A.

Moger, J. R. An exploratory study of questioning in the instructional processes in selected elementary schools. (Doctoral dissertation, Columbia University) Ann Arbor, Michigan: University Microfilms, 1966. No. 66-2661.

Nie, N. B., Dale, H. and Hull, C. H. SPSS: Statistical Package for the Social Sciences. New York: McGraw-Hill Book Company, 1970.

Peterson, L. R. Short-term memory, Scientific American, 1966, 215, 90-95.

Prehm, H. J., and Crosson, J. E. The mentally retarded. Review of Educational research, 1969, 39, 5-24.

Robinson, H. B., and Robinson, N. M. The mentally retarded child. New York: McGraw-Hill, 1965.

Rosenshine, B., and Furst, N. Research in teacher performance criteria. In R. O. Smith (Ed.), Research in Teacher Education: A Symposium. New Jersey: Prentice-Hall, Inc., 1971. pp. 37-72.

Rothkopf, E. Z. Learning from written materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Rothkopf, E. Z. The concept of mathemagenic activites. Review of Educational Research, 1970, 40, 325-336.

Rothkopf, E. Z., and Bisbicos, E. E. Selective facilitative effects of interspersed questions on learning from written materials. Journal of Educational Psychology, 1967, 58, 56-61.

Rothkopf, E. Z., and Bloom, R. D. Effects of interpersonal interaction on the instructional value of adjunct questions in learning from written material. Journal of Educational Psychology, 1970, 61, 417-422.

Rouse, S. T. Effects of a training program on the productive thinking of educable mental retardates. American Journal of Mental Deficiency, 1965, 69, 66-73.

Schreiber, J. E. Teachers question-asking techniques in social studies. (Doctoral dissertation, University of Iowa) Ann Arbor, Michigan: University Microfilms, 1967. No. 67-9099.

Schroder, H. M., Driver, M. J., and Streufert, S. Human information processing. New York: Holt, Rinehart, and Winston, 1967.

Semmel, M. I. Arousal theory and vigilance behavior of educable mentally retarded and average children. American Journal of Mental Deficiency, 1965, 70, 38-47.

Semmel, M. I. Language behavior of mentally retarded and culturally disadvantaged children. In J. F. Magory and R. B. McIntyre (Eds.) Fifth Annual Distinguished Lectures in Special Education. Los Angeles: University of Southern California Press, 1967, pp. 31-47.

Semmel, M. I., Barritt, L. S., Bennett, S. W., and Perfetti, C. A. Grammatical analysis of word associations of educable mentally retarded and normal children. American Journal of Mental Deficiency, 1968, 72, 567-576.

Semmel, M. I., Barritt, L. S., and Bennett, S. W. Performance of EMR and nonretarded children on a modified Cloze task. American Journal of Mental Deficiency, 1970, 74, 681-688.

Semmel, M. I., and Bennett, S. W. Effects of linguistic structure and delay on memory span of EMR children. American Journal of Mental Deficiency, 1970, 74, 674-680.

Semmel, M. I., and Williams, J. Intentional and incidental learning in normal, borderline, and retarded children. American Educational Research Journal, 1968, 5, 233-238.

Sitko, M. C. Input organizational strategies of educable mentally retarded and normal boys in free recall verbal learning. Unpublished Ph.D. dissertation, The University of Michigan, 1970.

Sitko, M. C., and Semmel, M. I. Organizational strategies in free recall verbal learning of normal and retarded children. Technical Report Center for Innovation in Teaching the Handicapped, Indiana University, Bloomington, Indiana, 1971.

Smith, R. M. Creative thinking abilities of educable mentally handicapped children in the regular grades. American Journal of Mental Deficiency, 1967, 71, 571-575.

Solomon, D., Bezdek, W. E., and Rosenberg, L. Teaching styles and learning. Chicago: Center for the Study of Liberal Education for Adults. (ERIC ED: 026-556), 1963.

Spaulding, R. L. Achievement, creativity, and self-concept correlates of teacher-pupil transaction in elementary schools. Hempsted, New York: Hofstra University (U.S. Office of Education Cooperative Research Project No. 1352), 1965.

Spitz, H. H. The role of input organization in the learning and memory of mental retardates. Chapter in N. R. Ellis (Ed.), International Review of Research in Mental Retardation, Vol. 2. New York: Academic Press, 1966, pp. 29-56.

Stevens, R. The question as a measure of efficiency in instruction: A critical study of classroom practice. Teachers College Contribution to Education, 1912, No. 48.

Stevenson, H. W., and Swartz, J. D. Learning set in children as a function of intellectual level. Journal of Comparative and Physiological Psychology, 1958, 51, 755-757.

Taba, H. Teaching strategies and cognitive functioning in elementary school children. Cooperation Research Project No. 2404, San Francisco State College, February, 1966.

Thompson, G. R., and Bowers, N. C. Fourth grade achievement as related to creativity, intelligence and teaching style. Paper presented to the American Educational Research Association, Chicago, February, 1968.

Tidsall, W. J. Productive thinking in retarded children. Exceptional Children, 1962, 29, 36-41.

Travers, R. M. W., Van Wagenen, R. K., Haygood, D. H., and McCromick, M. Learning as a consequence of the learner's task involvement under different conditions of feedback. Journal of Educational Psychology, 1964, 55, 167-173.

Van Wagenen, R. K., and Travers, R. M. W. Learning under conditions of direct and vicarious reinforcement. Journal of Educational Psychology, 1963, 54, 356-362.

Veldman, D. J. Fortran Programming for the Behavioral Sciences. New York: Holt, Rinehart, and Winston, 1967.

Watts, G. H., and Anderson, R. C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.

Wright, C. J., and Nuthall, G. Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. American Educational Research Journal, 1970, 7, 477-491.

Zeaman, D., and House, B. J. The role of attention in retardate discrimination learning. Chapter in N. R. Ellis (Ed.), Handbook of Mental Deficiency. New York: McGraw-Hill, 1963, pp. 159-223.

Zigler, E. Research on personality structure in the retardate. Chapter in N. R. Ellis (Ed.), International Review of Research in Mental Retardation, Vol. 1. New York: Academic Press, 1966, pp. 77-108.

APPENDIX A

THE STORY

1. Leroy Clark watched Mr. Pumkin leave the house. Mr. Pumkin had rented a room from Leroy's parents a week before, but none of them had learned very much about him.

Every morning at 7 o'clock Mr. Pumkin went off to work. Leroy always knew what Mr. Pumkin would be wearing. His shoes were so shiny that you could see your face in them. He wore a blue suit and a grey hat. Mr. Pumkin always wore the same thing.

No one knew where Mr. Pumkin worked. He would only say that he worked with money. He was such a well-dressed man that it looked as if he could buy his own house. Everyone thought it was strange that such a well-dressed man as Mr. Pumkin rented a room in Leroy's part of town.

Leroy watched as Mr. Pumkin went down the front steps to the big, black car parked in the street. No one in Leroy's neighborhood had ever owned such a big, new car.

Mr. Pumkin stopped beside his car. Leroy saw him reach far down in his pocket and take out his keys. And, at the same time, he saw something small and yellow fall out of Mr. Pumkin's pocket.

2. But Mr. Pumkin didn't seem to know that he had dropped something. He opened the car door, got in, started the engine and waved good-bye through the open window.

Leroy hurried from the window to the front door. "Mr. Pumkin! Mr. Pumkin!" he shouted. But Mr. Pumkin couldn't hear him.

Leroy ran to the street and looked down at the small piece of paper at his feet. When he picked it up, he saw that it was a large piece of paper that had been folded three times.

He looked down the street. Mr. Pumkin's car had gone away. He knew he shouldn't unfold the paper, even though he wanted to know what was inside. Mr. Pumkin had not given him permission to look at it. Leroy thought about taking the paper to Mr. Pumkin's room on the third floor or leaving the note at the foot of the stairs. But then Mr. Pumkin might not see it, or Leroy's mother might throw it away when she cleaned the house.

All at once, Leroy unfolded the paper. He knew it was wrong, but he had to know what was inside. He looked down at the paper, surprised by what he saw. It wasn't a letter at all. It was a map.

3. Leroy looked at the map for a long time. He followed each line with his finger, starting with the word "house." Next came "chimney," followed by "yard," "post," and "tree." There were five words altogether. Leroy wondered what it all meant.

He had read stories about buried treasure. Could Mr. Pumkin have a treasure buried?

APPENDIX A (cont'd)

Leroy went into the living room. He decided to tell his mother as soon as she came back from shopping. When he turned on the radio he heard a man giving the news. Leroy started to turn the radio off, but he heard something that made him stop.

The man was saying, "Police are still looking for the two well-dressed men who were seen driving away from the First Western Bank two days ago in a new, black car. The men are suspected of having stolen $50,000 from the bank."

After the news was over, Leroy started to think about Mr. Pumkin. He could still hear the words, "Two well-dressed men . . driving away in a new, black car. . . ."

Mr. Pumkin was well-dressed. And Mr. Pumkin had a dark, new car. Leroy looked down again at the map. Suddenly, he felt cold all over. Could Mr. Pumkin be a bank robber? And did this map show where he buried the money?

4. This time, Leroy knew what he was going to do. He opened the door and ran down the street. Officer Dawson was standing in the road. He was directing traffic. "Stop!" Leroy cried. "I have something important to tell you!"

Officer Dawson waited until Leroy caught up with him. "Hello, Leroy," he said. "What's wrong?"

Leroy stopped, out of breath. Quickly, he told Officer Dawson his story. He handed the map to the policeman and pointed to the circle. "And that's where I think he buried the money," Leroy explained. "Only I don't know where it is."

The policeman looked at the map and smiled. "Now, Leroy," he said. "Police don't put people in jail until they have proof that they committed a crime. This map doesn't mean anything by itself. We would have to find the money before we could take Mr. Pumkin to the jail over on Walnut Street."

"But he'll get away if you wait!" Leroy said.

"All right," said Leroy. "I'll find the money all by myself. I know he buried it."

Officer Dawson smiled again. "You be careful, young man," he said. "If your Mr. Pumkin is a robber, he's no man for you to tangle with. Boys can get into trouble when they try to be detectives. Leave that to the police."

5. Leroy knew the policeman didn't believe his story. Well, he thought to himself, he'd just show Officer Dawson that Leroy Clark could be a good detective!

Already Leroy thought he knew where the money was buried. He hurried back to his house and went to the center of his back yard. He looked around the yard, but all he could see was the old board fence. Then he saw something else!

Leroy saw the new fence post his father had put up at the corner of the yard a week ago. The light, fresh wood of the post stood out against the dark, old wood of the fence. And there was another thing. Over the top of the houses he could see one tree. It was the only tree he could see from his yard.

APPENDIX A (cont'd)

Leroy moved over until the tree stood behind the post. Leroy thought they were the tree and the post he had seen on the map. Then he looked at the hwo houses on the left side of his yard. The house in back of his own had a chimney. Not far away stood another tall house with a television antenna. Leroy was excited now.

6. Leroy moved over until he could see the television antenna above the chimney of the house. He turned to see if the fence post and the tree were still in a straight line, as the map showed. They were!

Leroy got down on the ground and began to crawl through the grass. The ground felt soft. He pulled at the grass, and a big piece came up in his hands. And he was sure that someone had been digging under the grass!

Jumping to his feet, Leroy ran down the street and caught up with Officer Dawson. "I've found it!" he cried. "I know where the money is buried!"

Officer Dawson frowned. "All right," he said. "Show me where the money is."

Leroy led the policeman back to his yard. He showed the officer the spot he had found.

Officer Dawson looked surprised. "Leroy," he said, "I thought you were just playing a game before. But now, I'm not so sure."

"What can we do?" Leroy asked.

"First of all, we should do some digging, too," the policeman said. "Do you have a shovel?"

"No," answered Leroy, "but I can borrow one from the people who live next door." Quickly, he brought the shovel back.

Just as the policeman began to dig, Leroy heard the back door bang.

7. Officer Dawson turned around and saw Mr. Pumkin and another man. At the same time, he whipped out his gun. "Get your hands in the air, you two!" he said.

Mr. Pumkin and the other man looked at the policeman in surprise. But they raised their hands. "What's this all about?" Mr. Pumkin said.

The Officer said, "We have reason to think you and your friend robbed a bank and buried the money in this yard. Is that true?"

Mr. Pumkin laughed. "No, that's not true, Officer," he said. "I did some digging here, but not to hide money."

"Oh?" said Officer Dawson. "Then why were you digging at all? This yard doesn't belong to you."

"I'll tell you, of course," Mr. Pumkin said. "But I must ask you to keep it a secret."

"We'll see about that after we hear what you have to say," Officer Dawson said.

Mr. Pumkin began to explain. "Well, I work for the Super-Fine Oil Company. This is my friend, Mr. Walker, who works for the company, too." He nodded toward the well-dressed man beside him.

APPENDIX A (cont'd)

"As you know," Mr. Pumkin went on, "there are oil wells in some parts of the city. And more oil wells have been found in Leroy's part of the city than in any other. There might be oil under this land, and we think it might be right under the top of the ground."

8. Officer Dawson asked Mr. Pumkin to tell him more about what he was doing. Mr. Pumkin said that his company had sent him to find out if oil was near the top of the ground here. "I rented a room at Leroy's house," he said, "and dug up the ground in his back yard. I had to test the dirt for oil."

"If what you say is true, why did you keep it a secret?" the policeman asked.

"That's easy," said Mr. Pumkin. "Another oil company called the Big-Money Oil Company is also interested in this land. And we didn't want them to know that we were making tests of the ground."

"So you're oil men," Officer Dawson said, "Can you prove it?"

"Of course," Mr. Pumkin answered.

Officer Dawson let the two men put their hands down. They reached into their pants pockets, took out their papers and gave them to the policeman. Then he made them put their hands back up while he read through the papers.

At last he said, "You can put your hands down now. I'm sure you work for an oil company, as you said. I'm very sorry about this, but Leroy found a map. . . ."

9. "So that's where my map went," Mr. Pumpkin said, sounding very surprised. "I missed it when I went to give it to Mr. Walker."

"I'm sorry, Mr. Pumkin," Leroy said. He looked down at his feet. He had never felt to ashamed. He knew he shouldn't have looked at the map in the first place.

"Well, I'll be going now," Officer Dawson said. "Leroy, you turned out to be a good detective, after all. But the next time you get on the trail of robbers, don't let me know about it. Okay?"

Leroy smiled. "I'm sorry," he said again. Officer Dawson then left to go back to directing traffic on the corner.

"Leroy, what made you think I was a bank robber?" Mr. Pumkin asked.

Leroy told the man about the news story he had heard on the radio. "I didn't really believe you were a bank robber, Mr. Pumkin. It was just that the map and everything made it look that way."

"Don't worry, Leroy," Mr. Pumkin said. "You did the right thing. I might have been a robber, for all you knew. But now I have some good news for you. The tests just came back. Your land does have oil under it. I think everyone around you will probably make quite a lot of money."

"Us, too?" Leroy asked.

"Oh, yes!" Mr. Pumkin said. "We will probably put an oil well right here in your back yard. Your family will get even more money than the others."

APPENDIX A (cont'd)

10. That night, Leroy's mother gave a party. They all had ice cream and cookies. Mr. Pumkin and Mr. Walker seemed very pleased at the way everything had turned out. And so were Leroy's parents.

Leroy was happy, too. He laughed when the men told how Leroy had thought they were bank robbers. People sometimes laugh at themselves when they are so wrong.

"I was right about the map," Leroy said. "It really was a treasure map."

Mr. Pumkin laughed. "It really was, Leroy. And it took a good detective to read it."

Then Mr. Pumkin and Mr. Walker got ready to go. As Mr. Pumkin was shaking Leroy's hand, Leroy saw a small piece of paper fall to the floor. "You dropped something," he said.

"I did?" Mr. Pumkin asked. "Well, just wait until I'm gone, and then you can see what it is."

After the men had gone, Leroy picked up the paper. He unfolded it and saw that it was a map. But this map showed a picture of his house, and there was a large X by the back door.

Leroy hurried to the back steps. By the door, he found a large package, which looked like a birthday present. It was covered with pretty paper, and there was a big bow on top. Quickly, he took off the bow and the paper. Inside the package was a large box. And on the box were the words SUPER-DETECTIVE SET.

APPENDIX B

REMEMBERING CRITERION TEST ITEMS

1. When was the first time that Mr. Pumkin missed the map?
2. What did Leroy hear the man say when he listened to him on the radio?
3. What did Leroy do after he got down on the ground?
4. What floor of Leroy's house did Mr. Pumkin live on?
5. How many times had the paper which Leroy found been folded?
6. What was the name of the other oil company which Mr. Pumkin didn't want to know where he was digging?
7. Where did Mr. Pumkin and his friend take out their papers from when Officer Dawson asked for them?
8. Where was Officer Dawson going when he left Leroy and Mr. Pumkin in the back yard?
9. How many words were on the map that Leroy found?
10. What was the name of the street that the jail was on?
11. What color was the "thing" the Leroy saw fall out of Mr. Pumkin's pocket?
12. What time did Mr. Pumkin go off to work?
13. Where did Leroy go to get a shovel for Officer Dawson?
14. What did Officer Dawson do when he saw Mr. Pumkin and the other man?
15. When Leroy got home after talking to Officer Dawson, where did he go?
16. Where was the present that Mr. Pumkin left for Leroy?
17. Where was Officer Dawson when Leroy caught up with him?
18. How many trees could Leroy see from his yard?
19. Who was Mr. Walker?
20. While at the party, Leroy said he was right about one thing. What was it he was right about?

APPENDIX B (cont'd)

INFERRING CRITERION TEST ITEMS

1. Why did Leroy feel ashamed?
2. Why did Leroy suspect Mr. Pumkin of being a bank robber?
3. How did Leroy know that someone had been digging under the grass?
4. Why did Leroy think that he shouldn't unfold the piece of paper?
5. Why couldn't Mr. Pumkin hear Leroy call to him?
6. Why did Mr. Pumkin rent a room in Leroy's house?
7. Why did the men have to give their papers to the policeman?
8. Why was Leroy's family going to get more money that the other families in the neighborhood?
9. What room of the house was the radio in?
10. Why did Officer Dawson want Leroy to be careful?
11. How did Leroy know what Mr. Pumkin would be wearing each morning?
12. What did people think it was strange for Mr. Pumkin to rent a room in Leroy's part of town?
13. Why did Officer Dawson want to do some digging in the spot that Leroy showed him?
14. Why did Officer Dawson take out his gun and tell the two men to get their hands up?
15. Where did Leroy think the money was buried?
16. How did Leroy go about finding where the present was located?
17. Why did Officer Dawson say they would have to find the money before they could do anything?
18. Why was Leroy so excited after he had looked all around his back yard?
19. Why did Mr. Pumkin think there was oil in Leroy's back yard?
20. Why did Leroy laugh when the men told of how Leroy thought they were bank robbers?

APPENDIX C

PRE-REMEMBERING TRAINING QUESTIONS

1. When you listen to the next part of the story, try to find out what color the "thing" was that Leroy saw fall out of Mr. Pumkin"s pocket.

2. When you listen to the next part of the story, try to find out how many times the paper which Leroy found had been folded.

3. When you listen to the next part of the story, try to find out how may words were on the map that Leroy found.

4. When you listen to the next part of the story, try to find out where Officer Dawson was when Leroy caught up with him.

5. When you listen to the next part of the story, try to find out how many trees Leroy could see from his yard.

6. When you listen to the next part of the story, try to find out where Leroy went to get a shovel for Officer Dawson.

7. When you listen to the next part of the story, try to find out what Officer Dawson did when he saw Mr. Pumkin and the other man.

8. When you listen to the next part of the story, try to find out where Mr. Pumkin and his friend took their papers out from when Officer Dawson asked for them.

9. When you listen to the next part of the story, try to find out where Officer Dawson was going when he left Leroy and Mr. Pumkin in the back yard.

10. When you listen to the next part of the story, try to find out what Leroy thought he was right about when he was at the party and said he was right about "one thing."

POST-REMEMBERING TRAINING QUESTIONS

1. What color was the "thing" that Leroy saw fall out of Mr. Pumkin's pocket?

2. How many times had the paper which Leroy found been folded?

3. How many words were on the map that Leroy found?

4. Where was Officer Dawson when Leroy caught up with him?

APPENDIX C (cont'd)

5. How many trees could Leroy see from his yard?

6. Where did Leroy go to get a shovel for Officer Dawson?

7. What did Officer Dawson do when he saw Mr. Pumkin and the other man?

8. Where did Mr. Pumkin and his friend take out their papers from when Officer Dawson asked for them?

9. Where was Officer Dawson going when he left Leroy and Mr. Pumkin in the back yard?

10. While at the party, Leroy said he was right about one thing. What was it he was right about?

PRE-INFERRING TRAINING QUESTIONS

1. When you listen to the next part of the story, try to find out why people thought it was strange for Mr. Pumkin to rent a room in Leroy's part of town.

2. When you listen to the next part of the story, try to find out why Mr. Pumkin couldn't hear Leroy call to him.

3. When you listen to the next part of the story, try to find out what room of the house the radio was in.

4. When you listen to the next part of the story, try to find out why Officer Dawson said they would have to find the money before they could do anything.

5. When you listen to the next part of the story, try to find out why Leroy was so excited after he looked all around his back yard.

6. When you listen to the next part of the story, try to find out why Officer Dawson wanted to do some digging in the spot that Leroy showed him.

7. When you listen to the next part of the story, try to find out why Officer Dawson took out his gun and told the two men to get their hands up.

8. When you listen to the next part of the story, try to find out why the men gave their papers to the policeman.

9. When you listen to the next part of the story, try to find out why Leroy's family was going to get more money than the other families in the neighborhood.

APPENDIX C (cont'd)

10. When you listen to the next part of the story, try to find out why Leroy laughed when the men told of how Leroy thought they were bank robbers.

POST-INFERRING TRAINING QUESTIONS

1. Why did people think it was strange for Mr. Pumkin to rent a room in Leroy's part of town?

2. Why couldn't Mr. Pumkin hear Leroy call to him?

3. What room of the house was the radio in?

4. Why did Officer Dawson say they would have to find the money before they could do anything?

5. Why was Leroy so excited after he had looked all around his back yard?

6. Wh did Officer Dawson want to do some digging in the spot that Leroy showed him?

7. Why did Officer Dawson take out his gun and tell the two men to get their hands up?

8. Why did the men have to give their papers to the policeman?

9. Why was Leroy's family going to get more money than the other families in the neighborhood?

10. Why did Leroy laugh when the men told of how Leroy thought they were bank robbers?

APPENDIX D

DIRECTIONS TO EXPERIMENTERS

1. Enter designated school.

2. Check in with principal.

3. Locate experimental room.

4. Locate subject's classroom.

5. Set up equipment.

6. Get subject from class; if he/she is unavailable take one of the other subjects you will do that day.

7. Ask subject to come with you. Tell him that you would like him to listen to a story and that you will also play some games.

8. Engage in rapport-building as you walk to experimental room.

9. Once in experimental room, situate subject so that you are facing each other with tape recorders off to the side.

EXPERIMENTAL PROCEDURES

10. Turn on the tape recorder and RECORD.

11. Tell subject that you would like him to listen to a mystery story which you will play on the tape recorder. Tell him also that the story is about a young boy who lives in a city and the problems he has when he tries to catch some bank robbers. Tell him also to expect a surprise ending.
Also tell him that it is important that he listen very carefully to the story and that you are going to ask him some questions about the story as he listens to it.
    11(a) In condition where questions are asked BEFORE each paragraph, be sure to tell subject that you will ask him a question before particular parts of the story. READ EXAMPLE TO SUBJECT.
    11(b) In condition where questions are asked AFTER each paragraph, be sure to tell subject that you will ask him a question after particular parts of the story. READ EXAMPLE TO SUBJECT.
    11(c) In condition where there are NO QUESTIONS just tell subject that every once in a while you will stop the story and take a short rest. Rest should be 20 seconds. If subject wants to talk during rest, it's okay to talk with him. But be sure not to lose track of time and DO NOT talk about story.

APPENDIX D (cont'd)

12. Turn on tape recorder with story and begin. Be sure to stop story when asking subject the question. Then start story again.

13. After subject gives answer to question, you may say THANK YOU or O.K. GIVE NO FEEDBACK! If subject can give NO ANSWER or says he forgot the question do not repeat question but say "That's okay, but try to remember the question the next time."

14. After having listened to all 10 sections of the story, you will take a 15 minute break during which time you should fill out information at top of page with the questions on it which you have just asked. Turn off tape while you play game, but be sure to turn it on again when you begin 20-item criterion test.

15. During break do not talk to subject about story. Tell him that you will be taking a short break and after that you will want to ask him some more questions about the story he just heard.

GAMES

16. Each experimenter will have a set of CUISENAIRE RODS. These can be used to make colorful designs, stick figures, to build houses or to teach math. Feel free to use them in whatever way you feel most comfortable. Here are some suggestions:

    (a) Make a stick figure with the rods and ask the subject to copy it.
    (b) Ask subject to make up his own figure.
    (c) Make up design and ask subject to copy.
    (d) Show subject the numerical relationship between rods and ask him to make up "10" as many ways as possible.

17. Be sure to keep aware of the time: 15 minutes.

18. You will now want to administer the 20-item test.

19. Each item should be read exactly as it is written on the test.

20. Each item may be read no more than 2 times in total. If subject says "I don't know" to an item, you should re-ask those questions once more after having gone through test once completely. Be sure to write down "I don't know" first, though.

21. Feedback to these questions can be accepting (e.g., O.K., Thank you, Uh huh) or positive (e.g., good, fine, etc.). The objective here is to keep the subject on task, happy and to prevent him from becoming anxious and up-tight.

22. If subject asks "Am I right?" or "Was that the right answer," etc., say you will tell him after you are finished asking all of the questions.

APPENDIX D (cont'd)

23. After the test is over you may discuss story with subject and answer his questions.

AFTER EXPERIMENT

24. Escort subject back to classroom.

25. Thank him VERY MUCH for cooperating and helping you.

26. Be sure to request that he not discuss the story with his friends because you want them to be surprised also when they hear the story. Request that he keep it a SECRET.

27. Pick up next subject.

APPENDIX E

ITEM ANALYSIS: REMEMBERING ITEMS[1]

ITEM NUMBER Relevant	ITEM NUMBER Incidental	R	D
	1	.45*G	35.56
	2	.40*G	40.00
	3	.56*G	40.00
	4	.49	22.22
5		.50*G	35.56
	6	.29	28.89
7		.60*G	48.89
8		.59*G	48.89
9		.43*G	28.89
	10	.45*G	44.44
11		.07	17.78
	12	.27	13.33
13		.16	66.67
14		.21	75.56
	15	.53*G	48.89
	16	.53*G	42.22
17		.46*G	28.89
18		.54*G	37.78
	19	.60*	17.78
20		.12*	20.00

[1]For purposes of item analysis, items had to be scored either zero or two. Items which had been scored half credit were thus given two points for purposes of item analysis only.

Kuder Richardson Reliability = .77

Standard Error of Measurement = 1.89

Spearman-Brown Reliability = .85

Standard Error of Measurement = 1.54

APPENDIX E (cont'd)

ITEM ANALYSIS: INFERRING ITEMS[1]

ITEM NUMBER Relevant	Incidental	R	D
	1	.47	13.33
	2	.39*G	71.11
	3	.30*G	55.56
	4	.34*G	31.11
5		.36*G	66.67
	6	.23	13.33
7		.34*G	71.11
8		.64*G	48.89
9		.40*G	33.33
	10	.13	8.89
	11	.65*G	53.33
12		.21	17.78
13		.35	86.67
14		.26	80.00
	15	.42	20.00
	16	.31*G	68.89
17		.61*G	42.22
18		.48*G	55.56
	19	.31*G	46.67
20		.61*G	37.78

[1]For purposes of item analysis, items had to be scored either zero or two. Items which had been scored half credit were thus given two points for purposes of item analysis only.

Kuder Richardson Reliability = .74

Standard Error of Measurement = 1.85

Spearman-Brown Reliability = .77

Standard Error of Measurement = 1.73

APPENDIX F

CORRELATION MATRIX OF ALL VARIABLES

Variables		1	2	3	4	5	6	7	8	9	10	11	12	13
SES of School	1	--												
Race of School	2	-.00	--											
Race of Child	3	-.21	.65	--										
Sex	4	-.06	-.19	-.16	--									
C.A.	5	.07	-.02	-.05	.00	--								
Question Type	6	.04	.01	.07	-.10	.14	--							
Question Position	7	.10	-.11	-.05	-.10	-.00	.64	--						
Criterion Type	8	-.04	.00	-.09	.05	-.06	.00	.00	--					
⌜	9	-.19	-.04	-.09	.02	.12	-.42	.16	-.16	--				
	10	-.07	-.00	-.00	.11	-.11	-.43	.31	.03	.30	--			
	11	-.04	-.12	-.08	.11	.04	-.13	-.30	-.01	.04	.21	--		
Between	12	-.16	.03	-.02	-.14	-.01	-.09	.21	-.03	.23	.24	.16	--	
Trials	13	-.15	-.00	.06	.15	-.01	-.35	-.14	.02	.11	.29	.51	.25	--
Items	14	.02	-.01	-.02	.08	-.03	-.44	.39	-.12	.31	.48	.12	.22	.19
by	15	-.03	.00	.03	-.02	-.01	.00	.29	-.14	.05	.30	.11	.24	.17
Section	16	-.09	.09	.09	-.11	.11	.00	.33	-.13	.18	.21	.01	.22	.08
⌞ From 1 to 10	17	-.13	.11	.06	-.10	.08	.12	.03	-.09	.10	.10	.45	.41	.46
	18	-.07	.00	-.11	-.09	-.07	-.12	.42	-.12	.36	.38	.15	.34	.14
Close Proximity Score	19	-.04	-.05	-.04	.03	.05	-.31	.46	-.12	.43	.58	.40	.50	.44
Distant Proximity Score	20	-.25	.07	.01	-.01	-.02	-.30	.09	-.13	.44	.57	.51	.55	.59
Total Between Trials	21	-.16	.00	-.02	.01	.00	-.34	.30	-.13	.48	.64	.51	.59	.57
Relevant	22	-.18	.03	.06	-.06	.05	.00	.04	-.31	.34	.31	.32	.32	.31
Incidental	23	-.10	.22	.14	-.31	.06	.17	.12	-.09	.22	.22	.33	.30	.28
Total	24	-.16	.14	.11	-.20	.07	.10	.09	-.24	.33	.30	.37	.35	.34
M.A.	25	-.16	.07	.02	-.06	.50	.08	-.01	-.02	.30	.05	.29	.23	.03
IQ	26	-.23	.10	.06	-.08	-.08	.00	-.02	.03	.25	.12	.31	.27	.05

$p < .05$ t value for statistically significant differences from zero with 90 dif. = .205

$p < .01$ t value for statistically significant differences from zero with 90 dif. = .267

APPENDIX F (cont'd)

CORRELATION MATRIX OF ALL VARIABLES

Variables		14	15	16	17	18	19	20	21	22	23	24	25	26
SES of School	1													
Race of School	2													
Race of Child	3													
Sex	4													
C.A.	5													
Question Type	6													
Question Position	7													
Criterion Type	8													
⌜	9													
	10													
	11													
Between	12													
Trials	13													
Items	14	--												
by	15	.33	--											
Section	16	.15	.33	--										
From 1 to 10	17	.09	.21	.35	--									
⌞	18	.18	.26	.41	.42	--								
Close Proximity Score	19	.57	.49	.55	.54	.60	--							
Distant Proximity Score	20	.43	.47	.37	.61	.56	.63	--						
Total Between Trials	21	.54	.53	.50	.63	.64	.90	.90	--					
Relevant	22	.07	.24	.33	.34	.32	.48	.45	.52	--				
Incidental	23	-.10	.15	.41	.40	.51	.40	.46	.48	.50	--			
Total	24	-.01	.22	.43	.42	.47	.51	.52	.58	.89	.84	--		
M.A.	25	.04	.28	.26	.21	.20	.27	.33	.34	.41	.40	.47	--	
IQ	26	.06	.34	.23	.20	.27	.27	.40	.38	.43	.43	.49	.82	--